

## METHOD FOR HYBRIDISATION OF IMMOBILIZED GENOMIC DNA

JCI.7. REC'D PCT/PTO 02 JUN 2005

**Field of the Invention**

- 5 The present invention relates to methods for intact genomic nucleic acid material hybridisations and detection and quantification of target nucleic acids in a genomic DNA sample.
- The present invention relates in particular to methods for automated multiple amplifiable probe hybridisations onto genomic DNA.
- 10 The methods of the present invention are particularly useful in screening methods for detection of copy number and changes in copy number of genomic DNA.

**Background of the Invention**

- Abnormalities of DNA copy number account for many genetic diseases in living
- 15 organisms, including many human genetic disorders. The largest of these abnormalities involve changes in copy number in entire chromosomes; for example in monosomies and trisomies (for example trisomy 21 resulting in Down syndrome), and segmental abnormalities such as 5p deletion in cri-du-chat syndrome. Alternatively, genetic diseases such as for example *DMD* (Duchenne muscular dystrophy), *BRCA1*
- 20 (breast cancer) or *MSH2/MLH1* (hereditary nonpolyposis colorectal cancer or HNPCC) may evolve from smaller copy number changes in genomic DNA which are too small to be detected by conventional cytogenetics. Further, at the level of individual genes, specific inherited diseases can result from deletions or duplications involving individual exons or entire genes.

- 25 The detection of changes in copy number in a complex genome is not straightforward. Commonly applied methods for diagnosing genetic diseases due to copy number alterations involve for example quantitative multiple PCR, Southern blotting and comparative genomic hybridisation. These techniques, however, albeit
- 30 commonly practiced and to a great extent recognized for their reliability are subject to a number of disadvantages. Southern blotting is time consuming and duplications may be difficult to be detected. A disadvantage of multiplex PCR for example is its restriction to the number of loci which can be analysed simultaneously. Although comparative genomic hybridisation can analyse a whole genome in a single test,
- 35 resolutions were proven to be relatively low. It is clear from the above that detection of copy number in a complex genome has a great technical challenge.

Despite its fundamental importance, it is only recently that systematic approaches have been developed to assess copy number at specific genetic loci, or to examine intact genomic DNA for sub-microscopic deletions of unknown location. New approaches include for example multiplex amplifiable probe hybridisation as  
5 described in WO 00/53804 (Armour). Although the power and specificity of multiplex amplifiable probe hybridisation is proven by the simultaneous assessment of copy number at large sets of human loci, this technique suffers from the general disadvantage that in particular the handling step with regard to removal of unbound probes is quite time-consuming.

10

Due to the exponential growth of research activity and diagnostic development, demand for improved hybridisation procedures is imperative and those skilled in the art recognize that it would be a distinct advantage in diagnostic research and extremely beneficial in commercial diagnostics if a highly efficient and economic  
15 diagnostic tool would be available.

Although flow-through hybridisation methods known in the art have become appreciated over the last years for their efficiency in performing numerous analysis technologies, such methods are restricted to the hybridisation of probes to four types  
20 of nucleic acid probes: large sections of DNA, small DNA (including cDNA), RNA, and peptide nucleic acids.

The present invention describes the principle of a unique flow-through hybridisation process for immobilized undigested or intact genomic DNA and a device for the said  
25 purpose whereby the hybridisation time as well as the amount of reagents used for hybridisation can be reduced by many folds. In particular, the present invention enables analysis of undigested or intact genomic DNA, thus not requiring time-consuming pre-hybridisation manipulation steps such as required in fragmentation-based procedures.

30

The present invention aims at providing improved methods for the quantitative detection of nucleic acids in a genomic sample with high resolution.

It is a further object of the present invention to provide methods for the quantitative  
35 detection of nucleic acids in a genomic sample with a much improved sensitivity level.

It is a further object of the present invention to provide methods for the quantitative detection of nucleic acids in a sample possessing an improved time-management.

It is a further object to provide devices, apparatuses and kits for carrying out said  
5 methods.

### Summary of the Invention

In order to accurately quantify nucleic acids within a genomic nucleic acids sample, the present invention provides a method for hybridisation of probes onto immobilized  
10 intact genomic DNA comprising the steps of (a) providing intact genomic DNA and denaturing said intact genomic DNA; (b) immobilizing said denatured intact genomic DNA onto a matrix, said matrix comprising pore sizes within a range of 0.6  $\mu\text{m}$  to 2  $\mu\text{m}$  including the outer limits; (c) providing a set of probes and passing said probes through said matrix under conditions favouring hybridisation of the probes to its  
15 complementary sequence in said intact genomic DNA; and (d) washing off non-hybridised probes through said matrix, leaving formed hybridised intact genomic DNA/probe complexes for further analysis.

The present invention provides methods for flow-through genomic hybridisation  
20 which are fast (high-speed), highly sensitive, highly specific and miniaturized.

The present invention allows much decreased analysis time by using flow-through hybridisation technology combined with the use of undigested or undigested or intact or non-manipulated genomic DNA. As exemplified within the present specification,  
25 non-routine experimentation led to the surprising finding that only matrices with specific parameters fulfil the requirements of passing said probes through said matrix to its complementary sequence in said intact genomic DNA while assuring the most favourable hybridisation kinetics.

A general outline of the hybridisation methods provided by the present invention is  
30 given in Figure 1.

### Detailed Description of the Invention

Before the present method and solutions used in the method are described, it is to be  
35 understood that this invention is not limited to particular methods, components, or solutions described, as such methods, components, and solutions may, of course, vary.

In the present specification and the appended claims, the singular forms "a", "an", and "the" include plural references unless the context clearly dictates otherwise.

- 5 It should also be understood that the terminology used herein is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

Accordingly, definitions should not be understood to limit the scope of the invention. Rather, they should be used to interpret the language of the description and, where  
10 appropriate, the language of the claims. These terms may also be understood more fully in the context of the description of the invention. If a term is included in the description or the claims that is not further defined within the present description, or that cannot be interpreted based on its context, then it should be construed to have the same meaning as it is understood by those skilled in the art.

- 15 The present invention relates in particular to screening for sequence copy number and to screening for changes in copy number of a plurality of nucleic acid sequences in a genome or genomic sample.

The terms "genome", "genomic content", "genomic sample", "genomic DNA",  
20 "genomic nucleic acid material" and "genetic material" are used interchangeably throughout the present specification and mean the nucleic acid molecules in an organism or cell that are the ultimate source of heritable genetic information of the organism. For most organisms, a genome consists primarily of chromosomal DNA, but it can also include plasmids, mitochondrial DNA, and so on. For some organisms,  
25 such as RNA viruses, a genome consists of RNA. As used within the present specification, genomic DNA is undigested or intact unless otherwise stated.

The terms 'undigested genomic DNA' and 'intact genomic DNA' are used interchangeably throughout the present specification.

- 30 By "nucleic acid" is meant DNA, RNA, or other related compositions of matter that may include substitution of similar moieties. For example, nucleic acids may include bases that are not found in DNA or RNA, including, but not limited to, xanthine, inosine, uracil in DNA, thymine in RNA, hypoxanthine, and so on. Nucleic acids may also include chemical modifications of phosphate or sugar moieties, which can be  
35 introduced to improve stability, resistance to enzymatic degradation, or some other useful property.

The loss or reduction in the normal number of copies of a genetic sequence (deletion) or the increase in copy number (amplification) are of widespread general importance. Such genetic alterations are known to underlie phenotype characteristics both somatic and germline. The demonstration of the site and nature of such genetic alteration is critical in the identification of the genes responsible and to the development of appropriate and effective treatments and therapies.

The present invention provides for methods to obtain genetic information from samples containing or suspected to have genomic content. It is medically and/or environmentally and/or socially important to identify genomic disorders. It will be well appreciated that also e.g. infectious organisms may be identified and quantified in such samples for optimal treatment of infections or contamination and for maintaining public health.

Methods according to the present invention are particularly designed for probe hybridisation onto immobilized genomic nucleic acid material. Hybridisation methods according to the present invention are characterized in that undigested or intact genomic contents are immobilized and subjected to flow-through probe hybridisation techniques.

20

Within this context, it is another object of the invention to provide for the use of a method according to the present specification and as described herein for intact genomic DNA hybridisation.

The immobilized genetic material within the present invention originates from a sample to be analysed for the presence/absence of any genomic abnormality.

By the term "genomic abnormality" is meant any deviation from a normal genomic content status. A genomic content status characterizes the condition or part thereof of a sample or the corresponding whole from which said genomic content was identified and quantified. A genomic abnormality may prevail through for example mutation(s) at the level of entire chromosomes including deletions and duplications, segmental abnormalities, genomic DNA deletions and duplications at the level of individual genes, involving individual exons or entire genes.

Abnormalities or irregularities involving endogenes as well as exogenes may lead to genomic abnormalities.

Accordingly, a genomic abnormality may equally prevail through the presence of exogenous nucleic acids such as by way of example and not limitation: naked autonomous replicating nucleic acids including for example plasmids and viroids; and embodied autonomous replicating nucleic acids such as pathogens, parasites, and contaminants. Said pathogen, parasite, and contaminant may be algae, archaea, bacteria; viruses; fungi including yeasts, molds and mycorrhizae; nematodes; protozoa and microsporidae.

A sample containing or suspected to have a genomic content may be biological material or any material comprising biological material from which nucleic acids may be prepared and analysed for the qualitative and quantitative presence of particular nucleic acid sequences. Genomic nucleic acid material to be used in the methods of the present invention may be within its sample format for direct analysis. However, a particular useful format is provided when the sample is subjected to some preparation prior to use in the analysis of the present invention. Said preparation may involve the removal of non-nucleic acid debris and suspension/dilution of the pure or isolated nucleic acid material in water or an appropriate buffer.

The genomic material may be isolated from virtually any sample. However, usually, the sample is a biological or a biochemical sample. The term "biological sample," as used herein, refers to a sample obtained from an organism or from components (e.g., cells) of an organism. The sample may be of any biological tissue or fluid. Frequently the sample will be a "clinical sample" which is a sample derived from a patient. Such samples include, but are not limited to, sputum, cerebrospinal fluid, blood, blood fractions such as serum including foetal serum (e.g., SFC) and plasma, blood cells (e.g., white cells), tissue or fine needle biopsy samples, urine, peritoneal fluid, and pleural fluid, or cells there from. Biological samples may also include sections of tissues.

Hybridisation methods according to the present invention use intact genomic DNA that is isolated. Methods for genomic DNA isolation from various samples are well known in the art.

Typically, intact genomic DNA is denatured prior to immobilization. DNA may be denatured by boiling or other methods as well known in the art.

The denatured DNA is subsequently immobilized within a matrix.

The term "matrix" refers to a material in which genetic material may be enclosed or embedded (as for study). The term matrix encompasses a wide range of potential substrates that can be used for the immobilization of intact genomic DNA.

Generally, the matrix can be composed of any material which will permit  
5 immobilization of intact genomic DNA or nucleic acids and which will not melt or otherwise substantially degrade under the conditions used to immobilize said genomic material and which allows hybridisation of said immobilized genomic material with probes by flow-through hybridisation.

A number of materials suitable for use as matrices in the present invention have  
10 been described in the art. Materials particularly suitable for use as matrices in the present invention include any type of permeable synthetic materials or natural materials provided that the pore diameter in case of a porous matrix or the mesh size in case of a matrix network allow for the permeation of the intact genomic nucleic acid material. Suitable matrix materials have pore sizes comprised within a range of  
15 0.6  $\mu\text{m}$  to 2.0  $\mu\text{m}$  including the outer limits; e.g. 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, and 2.0  $\mu\text{m}$ . Particular suitable pores sizes are comprised within a range of 0.6  $\mu\text{m}$  to 1.2  $\mu\text{m}$  including the outer limits; e.g. 0.6, 0.7, 0.8, 0.9, 1.0, 1.1 and 1.2  $\mu\text{m}$ . One of suitable materials as exemplified within the present specification is Whatman 3MM Chr paper. This exemplified material should however  
20 not be taken as in any way limiting. It will be well appreciated by a person skilled in the art that matrix thickness may vary in function of matrix strength, i.e. the thinner a matrix may be; the stronger the material is, wherein thin material might still support flow-through hybridisation and/or flow-through of reaction components. A particular suitable matrix is a thin pore matrix. Advantages, among others, of such particular  
25 suitable matrix is the limited amount of genomic DNA material needed. A particular suitable matrix thickness is within a range of 0.1 mm to 1 mm, including the outer limits. A more particular suitable matrix thickness is within a range of 0.3 mm to 0.5 mm including the outer limits.

30 It will be well appreciated that a combination of matrix materials may be envisaged in obtaining a desired matrix format characterized by a desired pore size(s) and strength.

Accordingly, in one embodiment of the present invention, a hybridisation method is  
35 provided, wherein said denatured intact genomic DNA is permeated within said matrix.

A matrix may be in the form of sheets, films or membranes and are permeable. For example, a matrix may consist of fibres such as glass wool or other glass or synthetic fibres such as plastic fibres, polyamide fibres (e.g. nylon) and the like. A matrix may equally consist of animal fibres such as silk and wool, or plant or vegetable fibres  
5 such as cotton, cellulose fibres and nitrocellulose fibres or cellulosic fibres including for example acetate and triacetate.

The matrix may be planar or have a simple or a complex shape. Particular useful matrices are membranes comprising a 3D network structure of which the surface to  
10 which the genomic nucleic acids are adhered is external surface as well as internal surface of the matrix. However, as will be appreciated in the art, it will be predominantly the internal surfaces that will have adhered thereto the genomic DNA.

Accordingly, in one embodiment of the present invention, a hybridisation method is  
15 provided, wherein the matrix is a membrane.

In a further embodiment of the present invention, a hybridisation method is provided, wherein said membrane comprises a 3D network structure.

20 In yet a further embodiment of the present invention, a hybridisation method is provided, wherein said network structure is a fibre network structure.

In yet a further embodiment of the present invention, a hybridisation method is provided, wherein said fibre is of vegetable origin.  
25

In yet a further embodiment of the present invention, a hybridisation method provided, wherein said fibre is cellulose.

The principle of the present invention is using a flow-through mechanism by which  
30 the probes pass through the membrane structure, allowing these probes coming in close contact with the corresponding complementary sequences within the immobilized genomic nucleic acids so that the target sequences can be effectively detected in high sensitivity and specificity.

35 Accordingly, in one embodiment of the present invention, a hybridisation method is provided, wherein said network structure is a flow-through structure.



Whatever the substrate or matrix material, there are a number of ways that nucleic acids can be attached or immobilized. Most common are physical adsorptive processes or chemical linking processes, including ultraviolet (UV) or covalent methods.

5

A great deal of interest has been shown in the production of activated membranes, that is, membranes that have direct reaction chemistry available on their surfaces. The ability to apply an aqueous sample directly to the membrane, and to have binding occur naturally without the need for further reagents, offers significant advantages, these are especially known with respect to the production of microarrays. Several different activation chemistries are well known in the art.

Accordingly, in one embodiment of the present invention, a hybridisation method is provided, wherein the matrix is activated with an affinity conjugate.

15

In a further embodiment, a hybridisation method is provided, wherein said affinity conjugate is chosen from the group comprising poly-L-lysine, poly-D-lysine, 3-aminopropyl-triethoxysilane, poly-arginine, polyethyleneimine, polyvinylamine, polyallylamine, tetraethylenepentamine, ethylenediamine, diethylenetriamine, triethylenetetramine, pentaethylenhexamine and hexamethylenediamine.

20

In yet a further embodiment, a hybridisation method is provided, wherein said affinity conjugate is poly-L-lysine.

Handling of the matrix is greatly improved by means of a device for holding said matrix such as described in PCT/EP02/02446 which is herewith incorporated by reference.

25

High-throughput analysis of numerous samples simultaneously may be accommodated by a system such as described in European Application No. 02076728.1, herewith incorporated by reference, which discloses a system for conducting bioassays, comprising a substrate plate with a number of wells, and an incubation device for holding the plate. This known analytical test device is composed of a plastic support wherein openings in the plastic support define wells with a certain diameter, said wells being open at the top for sample or probe application and having a substrate defining the bottom of each well. Said substrate may be a matrix as described in the present specification.

30

35

A system as described above allows for parallel processing of a large number of genomic nucleic acid samples and may be applied in automated robotic platforms. Such system usually comprises a microplate with an array of wells arranged in rows and columns, wherein the bottom of each well is a matrix having a flow-through fibre network. Using for example a microplate with an array of ninety-six wells allows a parallel processing of a large number of hybridisations resulting in a very efficient high-throughput analysis.

Accordingly, it is an object of the present invention to provide a device for flow-through hybridisation of probes onto immobilized intact genomic DNA comprising a well holder, said well holder comprising one or more round wells with a fixed diameter, said wells exposing a fibre network matrix, said matrix comprising pore sizes within a range of 0.6  $\mu\text{m}$  to 2  $\mu\text{m}$  including the outer limits; wherein said matrix permits immobilization of intact genomic DNA and which allows hybridisation of said immobilized intact genomic material with probes by flow-through hybridisation.

In one embodiment of the present invention, a device for flow-through hybridisation of probes onto immobilized genomic DNA is provided wherein said matrix permits permeation of intact genomic DNA.

Application of a pressure difference over the matrix will force the probes through the matrix 3D network structure, thereby creating a low pressure within the wells. By removing said low pressure, the probes are automatically forced back through the network of the matrix. Higher pressures will create a more rapid flow of the probes through the matrix structure. By alternatingly creating a low pressure over the matrix and removing the low pressure, the probes are forced through the matrix network a number of times.

Accordingly, it is another object of the present invention to provide an apparatus for flow-through hybridisation of probes onto immobilized intact genomic DNA comprising:

(a) a device for flow-through hybridisation of probes onto immobilized intact genomic DNA comprising a well holder, said well holder comprising one or more round wells with a fixed diameter, said wells exposing a fibre network matrix;

(b) means for addition of a controlled amount of fluid to at least one of the wells of the device as described in (a);

(c) means for applying and/or maintaining a controlled pressure difference over the matrix in each of the wells.

5

An optimum flow-rate is such that the residence time of the probe molecules near the immobilized target sequences is sufficient to generate hybridisation events in the shortest possible time.

Hybridisations are usually performed with flow rates comprised between  
10 50mm/30min and 250mm/30min including the outer limits. Particular suitable flow rates are comprised between 75mm/30min and 200mm/30min including the outer limits. More particular suitable flow rates are comprised between 100mm/30min and 150mm/30min including the outer limits. Usually, a particular suitable flow rate is 130mm/30min including the outer limits.

15

Accordingly, in one embodiment of the present invention, hybridisation methods are provided, wherein the matrix allows for a flow rate comprised between 50mm/30min and 250mm/30min including the outer limits.

20 A flow-through incubation, as employed in the methods as described herein, gives significantly reduced hybridisation times. Positive or negative pressure may be applied to the matrix in order to pump the probe solution dynamically up and down through the matrix pores or matrix network which may enhance the diffusion of the probes to the target sequences within the immobilized genomic material.

25

The duration of one cycle of forward and backward flow of probe solution across the matrix membrane may be comprised between 10 min and 1 sec. Usually, duration of one cycle is comprised between 5 min and 10 sec. More usually, duration of one cycle is comprised between 5 min and 10 sec. Yet more usually, duration of one  
30 cycle is comprised between 1 min and 45 sec. A particular suitable example of duration of a single cycle of forward and backward flow is 30 sec.

In another embodiment of the present invention, a hybridisation method is provided, wherein said probes are passed through said matrix by at least one cycle of  
35 alternating downwards and upwards flow.

It is common to perform analysis at a single constant temperature; the preferred temperature will depend on the envisaged hybridisation stringency. Adjustment of the hybridisation temperature may be accomplished by coupling of the matrix via a holding device to a heating device such as a water bath or a conductive heating plate. Alternatively, an incubator system with a temperature control system may be provided wherein said holding device comprising one or more matrices may be housed. Such incubator system is described in for example PCT/EP02/02448 which is hereby incorporated by reference.

Once sufficient time has elapsed or sufficient flow cycles have elapsed to provide for hybridisation, unbound probes are washed thoroughly away by means of a post-hybridisation flow-through wash step. A dynamical pumping allows immediate and highly efficient removal of any unbound probe in as little as one single downward flow. In general, flow conditions with regard to number of cycles and flow speed may be varied according to the envisaged stringency.

In one embodiment of the present invention, a hybridisation method is provided, wherein the washing step is carried out by passing through the matrix a wash fluid by at least one cycle of downwards flow.

Any bound probe is subsequently recovered and amplified. Said recovering and amplification may be essentially simultaneously, i.e. the probe-recovering step may be performed in a nucleic acid amplification environment. For example, nucleic acids immobilized onto a membrane and bound to identifier probes may be immersed in nucleic acid amplification buffer comprising amplification components. Setting of denaturing conditions will set free the bound identifier probes which then may be essentially simultaneously amplified.

Probe amplification involves the amplification (i.e. replication) of the identifier probe sequence being bound to the immobilized sample genomic nucleic acids, resulting in a significant increase in the number of identifier probe molecules.

Although numerous amplification techniques are known in the art, a particular suitable amplification technique employs a single primer pair, whereby each member of said primer pair is complementary to a primer binding sequence which is positioned flanking 5' or 3' to each identifier probe; said 5' and 3' flanking primer binding sequences being the same or substantially the same for each probe.

Accordingly, in one embodiment of the present invention, a method for hybridisation of probes onto immobilized intact genomic DNA is provided wherein the probes are flanked by primer binding sequences,

5

As used herein, "amplification" refers to the increase in the number of copies of a particular nucleic acid of interest wherein said copies are also called "amplicons" or "amplification products". In particular, by amplification is meant a technique for linearly or exponentially increasing the copy number of a nucleic acid molecule.

10

In one embodiment of the present invention, a hybridisation method is provided wherein the amplification of the recovered hybridised probes is a quantitative amplification.

15 The requirement for sensitivity (i.e. low detection limits) has been greatly alleviated by the development of the polymerase chain reaction (PCR) and other amplification technologies, which allow researchers to amplify exponentially a specific nucleic acid sequence before analysis. Suitable amplification methods include exponential amplification methods such as for example PCR, quantitative PCR (Q-PCR), biotin  
20 capture PCR as well as linear amplification methods such as for example linear amplification by *in vitro* transcription TYRAS and NASBA.

In a further embodiment of the present invention, a hybridisation method is provided wherein said amplification is by means of polymerase chain reaction.

25

The polymerase chain reaction (PCR) is widely used and described, and involves the use of primer extension combined with thermal cycling to amplify a target sequence; see US 4,683,195 and US 4,683,202, and PCR Essential Data, J. W. Wiley & sons, Ed. C. R. Newton, 1995, all of which are incorporated by reference.

30

Suitable PCR-based amplification strategies are well known in the art and may be, by way of example and not limitation, routine quantitative PCR (QC-PCR), reverse transcriptase PCR or RT-PCR, biotin capture PCR, nucleic acid sequence based amplification (NASBA), and TYRAS.

35

The TYRAS amplification method as disclosed in WO 99/43850, hereby incorporated by reference, is a non-selective poly-A mRNA amplification method which does not

encompass cDNA synthesis. The method comprises the hybridisation of an oligonucleotide, encompassing an oligo-T stretch, to the poly-A tail of the mRNA followed by RNase H digestion opposite the oligonucleotide and extension of the newly formed 3' end of the mRNA with reverse transcriptase. In this way the T7 RNA polymerase recognition sequence (i.e. T7 promoter) that is part of the oligonucleotide encompassing an oligo-T stretch is made double stranded. Upon binding of the T7 RNA polymerase to the promoter the original mRNA molecules are transcribed in multiple RNA copies of the opposite polarity.

RNA may also be amplified according to the method as disclosed in US Patent No 5,545,522 (Van Gelder), hereby incorporated by reference, wherein cDNA is synthesized from an RNA sequence using a complementary primer linked to an RNA polymerase promoter region complement and then antisense RNA (aRNA) is transcribed from the cDNA by introducing an RNA polymerase capable of binding to the promoter region.

Nucleic acid sequence based amplification (NASBA) is generally described in US 5,409,818 and "Profiting from Gene-based Diagnostics", CTB International Publishing Inc., N.J., 1996, both of which are incorporated by reference. NASBA relies on the simultaneous activity of 3 enzymes: AMV-RT (Avian Myoblastosis Virus-Reverse Transcriptase), RNase H and T7 RNA polymerase. NASBA is a special case of the 3SR amplification reaction or self-sustained sequence replication-reaction.

The 3SR reaction is a very efficient method for isothermal amplification of target DNA or RNA sequences in vitro. This method requires three enzymatic activities: reverse transcriptase, DNA-dependent RNA polymerase and *Escherichia coli* ribonuclease H.

For use in multiplex PCR, a primer should be designed so that its predicted hybridisation kinetics are similar to those of the other primers used in the same multiplex reaction. While the annealing temperatures and primer concentrations may be calculated to some degree, conditions generally have to be empirically determined for each multiplex reaction. Since the possibility of non-specific priming increases with each additional primer pair, conditions must be modified as necessary as individual primer sets are added. Moreover, artefacts that result from competition for resources (e.g. depletion of primers) are augmented in multiplex PCR, since differences in the yields of unequally amplified fragments are enhanced with each cycle.

As well known by the person skilled in the art, probe design for multiplex PCR may encompass flanking primer binding sequences that do not hybridise to the target sequence to overcome the above-mentioned draw-backs.

5

Accordingly, in one embodiment of the present invention, a method for target nucleic acid detection and quantification in an intact genomic DNA sample is provided, wherein each probe is flanked 5' and 3' by primer binding regions with said 5' and 3' flanking primer binding sequences being the same or substantially the same for each probe.

10

By target sequence or target nucleic acid is meant a nucleic acid sequence that a probe is designed to detect; e.g., for an "identification"-probe, the target sequence might be an identification sequence. By identification sequence is meant a nucleic acid sequence that is diagnostic of a particular organism or group of organisms or that is diagnostic for a particular genetic disease state when its presence or existence is assayed in a genome or enriched genome by hybridisation using the appropriate melting temperature criteria. By an enriched genome or enriched genomic fraction, is meant a genome or genomic fraction that has undergone an enrichment procedure that generates a selected fraction of the original genome. For example, for the purpose of genomic profiling, enriched genomes offer robust hybridisation-based diagnostics.

15

20

The set of identifier or identification probes or polynucleotides may correspond to particular mutations that are to be identified in a known sequence. As such, for a known gene that may contain any of several possible identified mutations, the set can comprise polynucleotides corresponding to the different possible mutations. This is, for instance, useful for genes like oncogenes and tumour suppressors, which frequently have a variety of known mutations in different positions.

30

In one embodiment of the present invention, a method for target nucleic acid detection and quantification in an intact genomic DNA sample is provided, wherein the amplified probes are provided with a label.

35

The term label as used in the present specification refers to a molecule propagating a signal to aid in detection and quantification. Said signal may be detected either visually (e.g., because it has a coloured product, or emits fluorescence) or by use of

a detector that detects properties of the reporter molecule (e.g., radioactivity, magnetic field, etc.). In the present specification, labels allow for the detection of the identification and quantification of target sequences within an intact genomic sample. Detectable labels suitable for use in the present invention include but are not limited to any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means.

Accordingly, virtually any label that produces a detectable, quantifiable signal and that is capable of being attached to a nucleotide and incorporated into the generated amplicon, can be used in conjunction with the methods of the invention. Suitable labels include, by way of example and not limitation, radioisotopes, fluorophores, chromophores, chemiluminescent moieties, chemical labelling such as ULS labelling (Universal Linkage system; Kretech) and ASAP (Accurate, Sensitive and Precise; Perkin Elmer), etc. Suitable labels may induce a colour reaction and/or may be capable of bio-, chemi- or photoluminescence.

Means for detecting such labels are well known to those of skill in the art. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted illumination. Enzymatic labels are typically detected by providing the enzyme with an enzyme substrate and detecting the reaction product produced by the action of the enzyme on the substrate; colorimetric labels are detected by simply visualizing the coloured label, and chemical labels by for example a platinum group form coordinative bonds with the labelling target, firmly coupling the label to the target.

Preferably, the position of the label will not interfere with generation, hybridisation, detection or other post-hybridisation modification of the labelled polynucleotide. A variety of different protocols may be used to generate the labelled nucleic acids, as is known in the art, where such methods typically rely on the labelled primers, or enzymatic generation of labelled nucleic acid using a labelled nucleotide. For instance, label may be incorporated into a nucleic acid during amplification steps in order to produce labelled amplicons. Alternatively, the generated amplicons may be labelled after the amplification.

A variety of labels may be employed, where such labels include fluorescent labels, isotopic labels, enzymatic labels, chemical labels, electron-dense reagents, particulate labels, etc. For example, suitable isotopic labels include radioactive



labels, e.g.  $^{32}\text{P}$ ,  $^{33}\text{P}$ ,  $^{35}\text{S}$ ,  $^3\text{H}$ ,  $^{125}\text{I}$ ,  $^{14}\text{C}$ . For example, suitable enzymatic labels include glucose oxidase, peroxidase, uricase, alkaline phosphatase etc. Other suitable labels include size particles that possess light scattering.

5     Fluorescent labels are particularly suitable because they provide very strong signals with low background. Fluorescent labels are also optically detectable at high resolution and quick scanning procedure. Fluorescent labels offer the additional advantage that irradiation of a fluorescent label with light can produce a plurality of emissions. Thus, a single label can provide for a plurality of measurable events.

10

Accordingly, in a further embodiment of the present invention, a method for target nucleic acid detection and quantification in an intact genomic DNA sample is provided, wherein the amplified probes are provided with a fluorescent label.

15     Desirably, fluorescent labels should absorb light above about 300 nm, usually above about 350 nm, and more usually above about 400 nm, usually emitting at wavelengths greater than about 10 nm higher than the wavelength of the light absorbed.

20     Particular useful fluorescent labels include, by way of example and not limitation, fluorescein isothiocyanate (FITC), rhodamine, malachite green, Oregon green, Texas Red, Congo red, SybrGreen, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein (JOE), 6-carboxy X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-  
25     carboxyfluorescein (5-FAM), N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA), cyanine dyes (e.g. Cy5, Cy3), BODIPY dyes (e.g. BODIPY 630/650, Alexa542, etc.), green fluorescent protein (GFP), blue fluorescent protein (BFP), yellow fluorescent protein (YFP), red fluorescent protein (RFP), and the like, (see, e.g., Molecular Probes, Eugene, Oregon, USA).

30

In one embodiment, the use of a method according to the present invention and as described herein is provided for detection and quantification of target nucleic acids in an intact genomic DNA sample.

35     Yet another object of the present invention is to provide for a method for target nucleic acid detection and quantification in an intact genomic DNA sample comprising the steps of: (a) providing intact genomic DNA and denaturing said

genomic DNA; (b) performing a hybridisation according to a method as described within the present specification; (c) recovering hybridised probes; and essentially simultaneously amplifying any recovered probe using a single primer pair, each member of said primer pair binding to each recovered probe onto the respective  
5 flanking primer binding sequences of said probe; and (d) qualitatively and quantitatively analysing the recovered amplified probes of step (c).

Recovered amplified probes may be analysed by gel electrophoresis. However, for improved reproducibility and accuracy of procedures, an automated system for  
10 determining genomic profiles is contemplated. In particular, the present invention connotes the use of a probe array or microarray which is interrogated with the amplified hybridised identifier probes provided by the methods of the invention. The term "probe array" relates to a substrate having a high density matrix pattern of positionally defined specific recognition reagents. The multiple probe copies provided  
15 by the method of the invention are capable of interacting, e.g. hybridising, with their specific counterparts, i.e. the specific recognition reagents, on the array. Because the specific recognition reagents are positionally defined, the sites of interaction will define the specificity of each interaction. The specific recognition reagents will typically be deoxyribonucleotide (DNA) probes, in which case said probe array is  
20 known as an oligonucleotide or cDNA array. Various array production methods are known in the art.

Accordingly, in a further embodiment of the present invention, a method for target nucleic acid detection and quantification in an intact genomic DNA sample is  
25 provided, wherein the analysis of the recovered amplified probes is by microarray analysis.

The term "nucleic acid" as used herein means a polymer composed of nucleotides, e.g. deoxyribonucleotides or ribonucleotides. The terms "deoxyribonucleic acid" and  
30 "DNA" as used herein means a polymer composed of deoxyribonucleotides. The term "oligonucleotide" as used herein denotes single stranded nucleotide multimers of from about 10 to about 100 nucleotides in length. The term "polynucleotide" as used herein refers to single or double stranded polymer composed of nucleotide monomers of from about 10 to about 5000 nucleotides in length, usually of greater  
35 than about 100 nucleotides in length up to about 1000 nucleotides in length.

For a given substrate size, the upper limit is determined only by the ability to create and detect the spots in the array. The preferred number of spots on an array generally depends on the particular use to which the array is to be put. For example, mutation detection may require only a small array. In general, arrays contain from 2  
5 to about  $10^6$  spots, or from about 4 to about  $10^5$  spots, or from about 8 to about  $10^4$  spots, or between about 10 and about 2000 spots, or from about 20 to about 200 spots.

Suitable arrays may be of any desired size, from two spots to  $10^6$  spots or even  
10 more. The upper and lower limits on the size of the substrate are determined solely by the practical considerations of working with extremely small or large substrates.

The immobilized polynucleotides may be as few as four, or as many as hundreds, or even more, nucleotides in length. Contemplated, as polynucleotides according to the  
15 invention are nucleic acids that are typically referred to in the art as oligonucleotides and also those referred to as nucleic acids. Thus, the arrays within the present invention are useful in applications where the generated identifier probe copies are hybridised to immobilized arrays of relatively short (such as, for example, having a length of approximately 6, 8, 10, 20, 40, 60, 80, or 100 nucleotides) detector probes.

20 The detector polynucleotides can be immobilized on the substrate using a wide variety of techniques. For example, the polynucleotides can be adsorbed or otherwise non-covalently associated with the substrate (for example, immobilization to nylon or nitrocellulose filters using standard techniques); they may be covalently  
25 attached to the substrate; or their association may be mediated by specific binding pairs, such as biotin and streptavidin.

A number of materials suitable for use as substrates for microarray analysis purposes in the instant invention have been described in the art. Exemplary suitable materials  
30 include, for example, acrylic, styrene-methyl methacrylate copolymers, ethylene/acrylic acid, acrylonitrile-butadienestyrene (ABS), ABS/polycarbonate, ABS/polysulfone, ABS/polyvinyl chloride, ethylene propylene, ethylene vinyl acetate (EVA), nitrocellulose, nylons (including nylon 6, nylon 6/6, nylon 6/6-6, nylon 6/9, nylon 6/10, nylon 6/12, nylon 11 and nylon 12), polycarylonitrile (PAN), polyacrylate,  
35 polycarbonate, polybutylene terephthalate (PBT), polyethylene terephthalate (PET), polyethylene (including low density, linear low density, high density, cross-linked and ultra-high molecular weight grades), polypropylene homopolymer, polypropylene

copolymers, polystyrene (including general purpose and high impact grades), polytetrafluoroethylene (PTFE), fluorinated ethylene-propylene (FEP), ethylene-tetrafluoroethylene (ETFE), perfluoroalkoxyethylene (PFA), polyvinyl fluoride (PVF), polyvinylidene fluoride (PVDF), polychlorotrifluoroethylene (PCTFE), polyethylene-chlorotrifluoro-ethylene (ECTFE), polyvinyl alcohol (PVA), silicon styreneacrylonitrile (SAN), styrene maleic anhydride (SMA), and glass.

Other exemplary suitable materials for use as substrates in the arrays of the present invention include metal oxides. Metal oxides provide a substrate having both a high channel density and a high porosity, allowing high density arrays comprising different specific recognition reagents per unit of the surface for sample application. In addition, metal oxides are highly transparent for visible light. Metal oxides are relatively cheap substrates that do not require the use of any typical microfabrication technology and, that offers an improved control over the liquid distribution over the surface of the substrate, such as electrochemically manufactured metal oxide membrane. Metal oxide membranes having through-going, oriented channels can be manufactured through electrochemical etching of a metal sheet. Metal oxides considered are, among others, oxides of tantalum, titanium, and aluminium, as well as alloys of two or more metal oxides and doped metal oxides and alloys containing metal oxides. The metal oxide membranes are transparent, especially if wet, which allows for assays using various optical techniques. Such membranes have oriented through-going channels with well-controlled diameter and useful chemical surface properties. Patent application EP-A-0 975 427 is exemplary in this respect, and is specifically incorporated in the present invention.

As will be appreciated in the art, methods according to the present invention find particular use in methods for genomic screening and gene expression studies. For example, intact genomic DNA or RNA can be immobilized onto a matrix as described herein and flow-through hybridised with for example two antisense oligonucleotides which leave a 10 base gap between them upon hybridisation onto a target sequence. By use of DNA polymerase in the presence of the necessary dNTPs, said gap will be filled which may then subsequently be ligated. Un-hybridised probes are then flow-through washed off through said matrix after which the hybridised oligonucleotides are eluted, quantitatively amplified and analysed by means of a microarray. As another example, intact genomic DNA or RNA can be immobilized onto a matrix as described herein and flow-through hybridised with for example short PCR-amplified probes.

It is another object of the invention to provide for the use of a method according to the present specification and as described herein for genomic screening.

5 By the term "genomic screening" is meant the screening for genetic variability within for example a genetic locus. Mutations may be located within genes for a variety of scenarios: e.g., for detecting sequence changes of HIV mutants which generate drug resistance and for detecting sequence changes of genes in relation to cancer development. Said sequence changes may include for example sequence deletions  
10 and sequence duplications.

Accordingly, in a further embodiment, the use of a method according to the present invention and as described herein is provided for detecting deletions or duplications in genomic DNA.

15

In a further embodiment, the use of a method according to the present invention and as described herein is provided for genome profiling.

By the term "genome profiling" is meant the identification of the presence and/or  
20 absence of genomic differences or variation between genomes of closely related species such as for example between humans and other primates. Genome profiling encompasses the identification of species using genotypes (genotyping).

In a further embodiment, the use of a method according to the present invention and  
25 as described herein is provided for identifying and quantitatively detecting the degree of pathogenesis, disease or contamination in a sample.

In a further embodiment, the use of a method according to the present invention and as described herein is provided for identifying and detecting the presence of  
30 infectious agents in a sample.

In a further embodiment, the use of a method according to the present invention and as described herein is provided for genotyping pathogens present in a sample.

35 The present invention also provides kits for performing the subject flow-through hybridisation methods. The subject kits at least include a device for flow-through hybridisation of probes onto immobilized intact genomic DNA comprising a well

holder, said well holder comprising one or more round wells with a fixed diameter, said wells exposing a fibre network matrix.

Accordingly, it is a further object of the present invention to provide a kit for flow-through hybridisation of probes onto immobilized intact genomic DNA comprising a device for flow-through hybridisation according as described in the present specification and instructions to carry out a method according to the specifications as described herein.

Kits may further comprise one or more reagents employed in the various methods, such as amplification primers for generating amplicons of the hybridised identifier probes as well as the amplification components. As used herein, the term "amplification components" refers to the reaction reagents such as enzymes, buffers, and nucleic acids including nucleotides necessary to perform an amplification reaction to form amplicons or amplification products of the hybridised identifier probes. A primer is a nucleic acid molecule with a 3' terminus that is either "blocked" and cannot be covalently linked to additional nucleic acids or that is not blocked and possesses a chemical group at the 3' terminus that will allow extension of the nucleic acid chain such as catalysed by a DNA polymerase or reverse transcriptase.

Accordingly, in one embodiment of the present invention a kit for target nucleic acid detection in an intact genomic DNA sample is provided additionally comprising (a) a set of probes, wherein each probe is flanked 5' and 3' by primer binding regions with said 5' and 3' flanking primer binding sequences being the same or substantially the same for each probe; (b) a single primer pair, each member of said pair being complementary to a primer binding region; (c) optionally amplification components allowing the amplification of any recovered hybridised probe; and (d) optionally a microarray, said microarray allowing analysis of the hybridisation results obtained by a method as described within the present specification.

### Short Description of the Figures

**Figure 1** shows a schematic representation of the hybridisation method of the present invention. Intact genomic material is first immobilized onto a suitable matrix such that said genomic material becomes permeated within said matrix material (step 1). A set of identifier probes is subsequently hybridised by flow-through onto the immobilized intact genomic material (step 2) to arrive at the formation of hybridised

intact genomic DNA/probe complexes (step 3). Unbound probes are washed off by flow-through washing (step 4), leaving said formed hybridised intact genomic DNA/probe complexes (step 5) for further analysis (step 6).

- 5 **Figure 2** illustrates the electrophoresis results of the PCR products corresponding to 12 control probes as prepared according to the description in Example 1, paragraph 1.1.1. Probe numbers, probe names, probe sizes and genomic locations are as mentioned in Table 1.

M, 100bp DNA ladder (Cat. No. 15628-050; Invitrogen).

10

**Figure 3** illustrates the electrophoresis results of the PCR products corresponding to the *MSH2* probes that were prepared from direct amplification of human genomic DNA using 18 specific primer pairs flanked by the same sequences as PZA and PZB as described in Example 1, paragraph 1.1.2. The sequences of the 18 specific *MSH2* primer pairs for PCR amplification are as mentioned in Table 2.

15

M, 100bp DNA ladder; Q, negative control.

**Figure 3A** illustrates said results after PCR in PCR mixture without DMSO;

**Figure 3B** illustrates said results after PCR in PCR mixture with DMSO.

- 20 **Figure 4** illustrates the electrophoresis results of the PCR products after flow-through hybridisation of 1 µg of immobilized intact genomic DNA from healthy control individual 1 with a control probe mixture in formamide hybridisation solution as described in Example 1, paragraph 1.3.

a, first PCR; b, repeat PCR reaction from the same multiplex amplifiable probe hybridisation solution; M, 100bp DNA ladder; P, PCR from control probe mix; Q, negative control for PCR; W, water control for hybridisation; 1, intact genomic DNA from healthy control individual 1. No PCR products were obtained from the negative controls from PCR and hybridisation.

- 30 **Figure 5** illustrates the electrophoresis results of the PCR products after flow-through hybridisation of 250ng of immobilized intact genomic DNA from three healthy control individuals (individuals 1, 4, and 5).

M, 100bp DNA ladder; P, PCR from control probe mix; Q, negative control for PCR; W, water control for hybridisation. No PCR products were obtained from the negative controls from PCR and hybridisation.

35

**Figure 6** illustrates the quality control of amplified probes as described in Examples 1 and 2.

**Figure 6A-C** illustrates gel-electrophoresis results of the quality control of 12 control probes, 18 *MSH2* probes and 19 *MLH1* probes.

5

**Figure 7** illustrates the quality of the PCR products after flow-through hybridisation of 1 µg of immobilized intact genomic DNA from healthy control individual 5 as described in Example 2.

M, 100bp DNA ladder; Q, negative control for PCR without hybridisation; W, PCR product obtained from water control after flow-through hybridisation and post-washes on Nylon membrane; 5, PCR product obtained from intact genomic DNA of individual 5 after flow-through hybridisation and post-washes on the Nylon membrane.

**Figure 8** illustrates the quality of the PCR products after flow-through hybridisation of 1 µg of immobilized intact genomic DNA from healthy control individual 5 as described in Example 3.

M, 100bp DNA ladder; Q, negative control for PCR without hybridisation; 5a, PCR product directly obtained from the intact genomic DNA of control individual 5 without hybridisation; W, PCR product obtained from water control after flow-through hybridisation and post-washes on Anodisc 25 membrane; 5b, PCR product obtained from the intact genomic DNA of control individual 5 after flow-through hybridisation and post-washes on Anodisc 25 membrane; P, PCR product directly obtained from PMPP probe mix without hybridisation.

25

## EXAMPLES

The following examples of the invention are exemplary and should not be taken as in any way limiting.

### 30 **Example 1: Flow-through hybridisation of intact genomic DNA on Whatman 3MM Chr paper**

#### 1.1 Probe preparation

##### 1.1.1 Probe preparation – control probes

35 Plasmids (100ng/µl) were obtained from the Institute of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK (Dr. John Armour). Probe DNA was amplified from these plasmids using flanking vector primers PZA



25

(AGTAACGGCCGCCAGTGTGCTG; SEQ ID No. 1) and PZB  
(CGAGCGGCCGCCAGTGTGATG; SEQ ID No. 2) (Isogene).

The PCR reaction was performed (PTC-200 Peltier Thermal Cycler; MJ Research  
5 INC; Massachusetts, USA) in a reaction mixture comprising 5µl 10x PCR gold buffer  
(PE) (Cat. No.4311816; Applied Biosystems), 2.5µl MgCl<sub>2</sub> (25mM), 1.25µl dNTPs  
(10mM; Amersham Pharmacia Biotech), 0.125µl AmpliTaq Gold® (5 U/µl) (Cat.  
No.4311816; Applied Biosystems), 1µl PZA forward primer (10pM), 1µl PZB reverse  
primer (10pM), 1µl plasmid DNA, and 38.125µl HLPC-water. A PCR program with  
10 following cycle order was completed: cycle 1, 3 min at 94°C; cycles 2 to 35, 1 min at  
94°C, 1 min at 60°C, 1 min at 72°C; and finally 10 min at 72°C.

Obtained PCR products were purified using the Qiaquick PCR purification kit (Cat.  
No.28106, QIAGEN, Germany) and dissolved in 300µl EB buffer (10mM Tris-Cl,  
pH8.5).

15

A 40-times dilution was made up from the purified PCR products (5µl PCR products +  
195µl water) for DNA concentration measurement using a SpectramaxPlus 384  
reader (Molecular Devices; Sunnyvale, CA, USA). The following concentrations were  
obtained: probe 1, 27ng/µl; probe 2, 26ng/µl; probe 3, 23ng/µl; probe 4, 11ng/µl;  
20 probe 5, 5ng/µl; probe 6, 12ng/µl; probe 7, 13ng/µl; probe 8, 13ng/µl; probe 9,  
19ng/µl; probe 10, 6ng/µl; probe 11, 22ng/µl and probe 12, 5ng/µl.

Subsequently, 10µl PCR products were loaded on a 2% agarose gel in 400mM Tris-  
Acetate/10mM EDTA (TAE) (0.5x; Cat. No. 15558-042; GIBCOBRL). Electrophoresis  
25 was performed at 100V for 40 min (Figure 2).

Probe names, sizes and sequences are given in Table 1.

#### 1.1.2 Probe preparation – MSH2 probes

Genomic DNA 1 was obtained from a healthy control individual via the blood bank of  
30 the University Hospital of Leiden, The Netherlands. *MSH2* probes were prepared by  
PCR on genomic DNA 1. The *MSH2* primers used are shown in Table 2.

PCR was performed in a reaction mixture comprising 5µl PCR Gold buffer (10X),  
2.5µl MgCl<sub>2</sub> (25mM), 1.25µl dNTPs (10mM), 0.125µl AmpliTaq Gold (5U/µl), 1µl  
*MSH2* forward primer (10pM), 1µl *MSH2* reverse primer (10pM), 1µl genomic DNA  
35 (100ng/µl), and 38.125µl HLPC-water.

Additionally, a second PCR mixture was made up including 5µl DMSO.

A PCR program with following cycle order was completed: cycle 1, 3 min at 94°C; cycles 2 to 5, 1 min at 94°C, 1 min at 56°C, 1 min at 72°C; cycles 6 to 10, 1 min at 94°C, 1 min at 53°C, 1 min at 72°C; cycles 11 to 35, 1 min at 94°C, 1 min at 50°C, 1 min at 72°C; and finally 10 min at 72°C.

5

Obtained PCR products were purified using the Qiaquick PCR purification kit and dissolved in 300µl EB buffer (10mM Tris-Cl, pH 8.5). A 40-times dilution was made up from the purified PCR products (5µl PCR products + 195µl water) for DNA concentration measurement using a SpectramaxPlus reader. Subsequently, 10µl  
10 PCR products were loaded on a 2% agarose gel in TAE (0.5X). Electrophoresis was performed at 100V for 40 min (Figure 3).

### **1.2 Poly-L-lysine (PLL) coating of Whatman 3MM Chr paper**

Whatman 3MM Chr paper (cat No. 3030 917) was cut into small pieces and 50 of  
15 them were placed in a Teflon holder (see WO 02/072268 for further specifications on said holder). A 0.01% poly-L-lysine solution was prepared with 35ml Poly-L-lysine (0.1%, Sigma; Cat. No. P 8920), 35 ml PBS (1X), and 280ml filtered HPLC water. This 350ml PLL solution was subsequently poured in a 600ml beaker which was placed on a plate shaker. The holder was gently moved up and down to prevent that  
20 air bubbles would be enclosed between the holder and the Whatman papers. The beaker was closed with parafilm and incubated on the plate shaker at room temperature with shaking at 100 rpm for 1 hour. Subsequently, Whatman papers were transferred to a second beaker filled with 350ml HPLC-water; the holder was again moved gently up and down for a couple of times. This transfer was repeated at  
25 least one more time. Papers were then transferred to an aluminium foil dish and placed for 2 hours in a vacuum oven at 37°C under vacuum. After turning off the vacuum pump, papers were allowed to cool down to room temperature after which they were stored in a dark and dry place.

### **1.3 Intact genomic DNA hybridisation on Whatman 3MM Chr paper coated with PLL whereby washing is performed with a reduced washing volume**

#### **1.3.3. Solutions**

50% deionised formamide hybridisation solution was prepared by mixing together 5ml formamide (100% deionised), 1.5ml SSC (20X), 0.5ml SDS (20%), 1ml  
35 Denhardtts solution (100X), 20µl Tween-20, and 1.8ml HPLC-water. Pre-hybridisation buffer was prepared by drying 10µl herring sperm DNA (10µg/µl) using a SpeedVac for 15min and adding 50µl of hybridisation buffer to this dried herring sperm DNA.

Probe mixture was prepared by mixing 4µl control probe mix (10ng/each probe), 20µl Cot-1 DNA, and 1ul blocker (50pmol/µl) of PZAX (AGTAACGGCCGCCAGTGTGCTGGAATTCTGCAGAT; SEQ ID No. 3) / PZBX (CGAGCGGCCGCCAGTGTGATGGA; SEQ ID No. 4), Cot-1 DNA (Cat. No.1581074; Roche) was used to block repetitive sequences of human genomic DNA for specific hybridisation. The probe mix was dried using a SC110A-240 SpeedVac Plus (Savant Instrument INC; New York, USA) for 15min after which 50µl of hybridisation buffer was added.

10 **1.3.2 Intact genomic DNA hybridisation on Whatman 3MM Chr paper coated with PLL**

A PLL-coated Whatman paper slide was made onto which 1µg denatured intact genomic DNA of control 1 was spotted. Water was spotted as control on an individual paper. The DNA was allowed to dry for 10 min at room temperature. The DNA was subsequently UV cross-linked (UV Stratalinker 1800; Cat. No.400072; Stratagene; California, USA) to the Whatman papers at 50 mJ on both sides. 50% deionised formamide hybridisation solution was prepared by mixing together 5ml formamide (100% deionised), 1.5ml SSC (20x), 0.5ml SDS (20%), 1ml Denhardt's solution (100x), 20µl Tween-20, and 1.8ml HPLC-water. The pre-hybridisation solution consisting of 100µg herring sperm DNA (10µg/µl; Cat. No.D1816; Promega) in hybridisation solution was boiled for 5 min and placed on ice. The Whatman papers were pre-hybridised in 50µl pre-hybridisation solution at 42°C for 2 hours using flow-through system (0.2 bar pressure at 42°C for 2 hours at 2 cycles/min). The pre-hybridisation solution was removed followed by a washing step once with hybridisation buffer.

The hybridisation probe mix was boiled for 5 min and placed on ice. 50µl hybridisation probe mix solution was added and incubated at 42°C for 4 hours using the flow-through system.

30

**1.3.3 Post-hybridisation washes**

The hybridisation mix was pipetted off and the Whatman papers were washed using flow-through system at 42°C using 25ml of solution 1 for 2.5 min and 25ml of solution 2 for 2.5 min. Wash solution 1 consisted of 1% SSC (20x SSC, 3M NaCl, 0.3M Sodium Citrate) and 1% SDS and wash solution 2 consisted of 0.1% SSC and 0.1% SDS.

35

Individual Whatman papers were transferred into 50µl of PCR buffer (1x) in 1.5ml tubes and each boiled for 5 min. 5µl of the boiled solution was transferred into a tube with PCR mixture comprising 5µl PCR gold buffer (10x), 2.5µl MgCl<sub>2</sub> (25mM), 1.25 µl dNTPs (10mM), 0.125 µl AmpliTaq Gold (5U/µl), 1µl PZA Forward primer (10pM), 1µl PZB reverse primer (10pM), 5µl sample solution, and 34.125µl HLPC-water.

A PCR program with following cycle order was completed: cycle 1, 3 min at 94°C; cycles 2 to 35, 1 min at 94°C, 1 min at 60°C, 2 min at 72°C; and finally 10 min at 72°C. 10µl PCR products were loaded onto a 2% agarose gel with TAE (0.5X). Electrophoresis was performed at 100V for 40 min. The results are shown in Figure 4.

In conclusion, the above experiment demonstrates that Multiplex Amplification Probe Hybridisation can be performed by flow-through hybridisation on intact genomic DNA immobilized within a matrix or membrane structure.

#### **1.4 Intact genomic DNA hybridisation on Whatman 3MM Chr paper coated with PLL using control probes**

##### **1.4.1 Solutions**

50% deionised formamide hybridisation solution was prepared by mixing together 5ml formamide (100% deionised), 1.5ml SSC (20x), 0.5ml SDS (20%), 1ml Denhardt's solution (100x), 20µl Tween-20, and 1.8ml HLPC-water. Pre-hybridisation buffer was prepared by drying 10µl herring sperm DNA (10µg/µl) using a SpeedVac for 15min and adding 30µl of hybridisation buffer to this dried herring sperm DNA. Probe mixture was prepared by mixing 4µl PMP22 control probe mix (10ng/each probe, see Table 1), 20µl Cot-1 DNA (1mg/ml), and 1µl PZAX/PZBX blocker (50pmol/µl). The end-blocking primers PZAX and PZBX as described in Example 1, paragraph 1.3 (Isogene) were added to prevent cross-hybridisation between different probes used in the same mixture. The probe mix was dried using a SpeedVac for 15min after which 30µl of hybridisation buffer was added.

##### **1.4.2 Intact genomic DNA hybridisation on Whatman 3MM Chr paper coated with PLL**

250ng of denatured intact genomic DNA of control individuals 1, 4 and 5 was spotted onto individual PLL-coated paper. Water was spotted as control on an individual paper.

The DNA was allowed to dry for 10 min at room temperature. The DNA was subsequently UV cross-linked to the Whatman papers at 50 mJ on both sides. The pre-hybridisation solution was boiled for 5 min and placed on ice. Whatman papers were pre-hybridised in 30µl pre-hybridisation solution at 42°C for 2 hours using flow-through system (at 0.2 bar and 2 cycles/min). The pre-hybridisation solution was removed followed by a washing step once with hybridisation buffer.

The hybridisation probe mix was boiled for 5 min and placed on ice. 30µl hybridisation probe mix solution was added and incubated at 42°C for 4 hours using the flow-through system.

#### 1.4.3 Post-hybridisation washes

Wash solutions were prepared as before.

The hybridisation mix was pipetted off and the Whatman papers were washed using flow-through system at 65°C using 50ml of solution1 and 50ml of solution 2 for 10 min. Individual Whatman papers were transferred into 50µl of PCR buffer (1x) in 1.5ml tubes and each boiled for 5 min. 5µl of the boiled solution was transferred into a tube with PCR mixture comprising 5µl PCR gold buffer (10x), 3µl MgCl<sub>2</sub> (25mM), 5µl dNTPs (2.5mM), 0.125µl AmpliTaq Gold (5U/µl), 1µl PZA Forward primer (10pM), 1µl PZB reverse primer (10pM), 5µl sample solution, and 22.375µl HLPC-water. A PCR program with following cycle order was completed: cycle 1, 3 min at 94°C; cycles 2 to 35, 1 min at 94°C, 1 min at 60°C, 2 min at 72°C; and finally 10 min at 72°C. 10µl PCR products were loaded onto a 2% agarose gel with TAE (0.5x). Electrophoresis was performed at 100V for 40 min. The results are shown in Figure 5.

In conclusion, the above experiment demonstrates that strong PCR bands were obtained upon electrophoresis of PCR products after flow-through hybridisation of a small amount of intact genomic DNA from individuals 1, 4 and 5 with the control probes.

#### **Example 2: Flow-through hybridisation of intact genomic DNA on Nylon membrane – 0.45 µm pore diameter**

In this experiment, Nylon membranes with 0.45 µm pore diameter (Amersham Biosciences, Cat No. RPN303B) were used to explore the use thereof in intact genomic DNA flow-through hybridisation.

## 2.1. Flow-through hybridisation of intact genomic DNA on Nylon membrane

### 2.1.1. Hybridisation

1  $\mu\text{g}$  of intact genomic DNA of control individual 5 and water were spotted on  
5 individual Nylon membranes. The DNA spot was dried for 10 minutes at room  
temperature. Subsequently the DNA was cross-linked to the filters at 50mJ on both  
sides of the membranes. 2 ml of pre-hybridisation solution was prepared by adding  
together 280  $\mu\text{l}$  1M  $\text{NaH}_2\text{PO}_4$ , 720  $\mu\text{l}$   $\text{Na}_2\text{HPO}_4$ , 700  $\mu\text{l}$  20% SDS, 276  $\mu\text{l}$  HLPC-water,  
20  $\mu\text{l}$  herring sperm DNA and 4  $\mu\text{l}$  of 0.5M EDTA. The Nylon filters were pre-  
10 hybridised in 50  $\mu\text{l}$  pre-hybridisation solution at 65°C for 2 hours using flow-through  
system as described in US 6,383,748 B1(0.2 bar pressure at 42°C for two hours at 2  
cycles/min). The pre-hybridisation solution was removed and replaced with 50 $\mu\text{l}$  of  
Cot-1 DNA solution and flown-through for 30 minutes at 65°C. The Cot-1 solution  
was subsequently removed and 50  $\mu\text{l}$  hybridisation solution comprising the PMP22  
15 control probes (see Table 1) was added. Flow-through hybridisation was performed  
at 65°C during 4 hours.

### 2.1.2. Post-hybridisation washes

For the post-hybridisation washes wash solutions 1 and 2 were prepared and  
20 incubated at 65°C. Wash solution 1 was prepared by diluting 25 ml 20% SSC and 25  
ml 20X SDS up to 500 ml with HLPC-water. Wash solution 2 was prepared by diluting  
2.5 ml 20% SSC and 2.5 ml 20X SDS up to 500 ml with HLPC-water. The  
hybridisation mixture was pipetted off from the Nylon membranes after which these  
membranes were washed subsequently with 50 ml wash solution 1 and 50 ml wash  
25 solution 2 at 65°C using flow-through system as described in US 6,383,748 B1.

### 2.1.3. PCR

The individual Nylon membranes were transferred into 50  $\mu\text{l}$  of HLPC-water in a 1.5  
ml tube and boiled for 5 minutes. 2  $\mu\text{l}$  of the boiled solution was then transferred into  
30 a new tube containing PCR reagents including 5  $\mu\text{l}$  10x PE buffer, 2.5  $\mu\text{l}$  25 mM  
 $\text{MgCl}_2$ , 1.25  $\mu\text{l}$  10mM dNTP, 0.125  $\mu\text{l}$  PE Taq (5 U/ $\mu\text{l}$ ), 1  $\mu\text{l}$  PZA forward primer  
(50pM; see also Example 1), 1  $\mu\text{l}$  PZB reverse primer (50pM; see also Example 1))  
and 37.125  $\mu\text{l}$  HLPC-water. A PCR program with following cycle order was  
completed: cycle 1, 5 min at 94°C; cycles 2 to 35, 45 sec at 94°C, 1 min at 57°C, 1  
35 min at 68°C; and finally 10 min at 68°C. 10 $\mu\text{l}$  PCR products were loaded onto a 2%  
agarose gel with TAE (0.5x). Electrophoresis was performed at 100V for 35 min. The

results are shown in Figure 7. Only weak PCR results from both water and individual 5 samples were obtained indicating unspecific and inefficient hybridisation.

In conclusion, example 2 shows that a 0.45 µm diameter pore sized membrane does  
5 not allow efficient flow-through of the hybridising probes through the porous membrane.

### **Example 3: Flow-through hybridisation of intact genomic DNA on Anodisc 25**

In this example, Multiple Amplification Probe Hybridisation (MAPH) was performed on  
10 intact genomic DNA immobilized onto Anodisc 25 membranes.

#### **3.1. Intact genomic DNA hybridisation on Anodisc 25 (0.2 µm pore size)**

##### **3.1.1. Genomic hybridisation**

Individual Anodisc 25 membranes silanised with 3-mercaptopropyltrimethoxysilane  
15 (MPS) were spotted with respectively 1 µg denatured intact genomic DNA from control individual 5 and water. DNA was cross-linked to the membrane by UV cross linking at 50mJ on both sides. Pre-hybridisation was carried out in 20 µl pre-hybridisation solution (see Example 3) at 65°C during 30 minutes using PamGene's flow-through system (0.2 bar pressure at 42°C for two hours at 2 cycles/min). The  
20 pre-hybridisation solution was removed and replaced with 20 µl of Cot-1 DNA solution and flown-through for 20 minutes at 65°C. The Cot-1 solution was removed and 20 µl hybridisation solution comprising PMP22 probes were added. Flow-through hybridisation was carried out at 65°C for 1 hour.

##### **3.1.2. Post-hybridisation washes**

For the post-hybridisation washes wash solutions 1 and 2 were prepared and incubated at 65°C. Wash solution 1 was prepared by diluting 25 ml 20% SSC and 25 ml 20X SDS up to 500 ml with HLPC-water. Wash solution 2 was prepared by diluting 2.5 ml 20% SSC and 2.5 ml 20X SDS up to 500 ml with HLPC-water. The  
30 hybridisation mixture was pipetted off from the Anodisc 25-MPS membranes after which these membranes were transferred into two 1.5 ml tubes, one for the DNA sample and one for the water control. Membranes were washed subsequently with wash solution 1 for 30 minutes and with wash solution 2 for 45 minutes at 65°C.

##### **3.1.3. PCR and results**

Individual washed membranes were transferred into 50 µl of 1x PCR buffer in 1.5 ml tubes and boiled for 5 minutes. 5 µl of the boiled solutions were subsequently

transferred into 0.5 PCR thin-wall tubes. The following PCR mixture was added: 5 µl 10x PE buffer (Perkin Elmer), 2.5 µl 25 mM MgCl<sub>2</sub>, 1.25 µl 10mM dNTP, 0.125 µl PE Taq (5U/µl), 1 µl PZA forward primer (50pM), 1 µl PZB reverse primer (50pM) and 34.125 µl HLPC-water. A PCR program with following cycle order was completed:

5 cycle 1, 5 min at 94°C; cycles 2 to 35, 45 sec at 94°C, 1 min at 57°C, 1 min at 68°C; and finally 10 min at 68°C. 5µl PCR products were loaded onto a 2% agarose gel with TAE (0.5x). Electrophoresis was performed at 100V for 45 min. The results are shown in Figure 8. Only weak PCR results from both water and individual 5 samples were obtained indicating unspecific and inefficient hybridisation. This is due to the

10 fact that that the small pore size did not allow the passing-through of the probes and an efficient post-hybridisation wash could not be established.

**Example 4: Flow-through hybridisation of intact genomic DNA on Anodisc 25 membrane – 0.2µm pore diameter (Whatman Plc.)**

15

In this experiment, Anodisc 25 membranes (Whatman) were first positively charged by silanation with 3-aminopropyltriethoxysilane (APS). The purpose of the experiment was to evaluate the 0.2 µm-pore-size-membranes for use for intact genomic DNA flow-through hybridisation. Prior to hybridisation, the silanised

20 membranes were blocked with either herring sperm DNA or with acetic anhydride and N,N-diisopropylethylamine.

**4.1. Materials and reagents**

1. Anodisc 25, 0.2 µm membrane (Whatman)
- 25 2. 3-aminopropyltriethoxysilane (Acros, APS)
3. Acetic anhydride, N,N-diisopropylethylamine, DMSO, Dioxane, Acetonitril and dichloromethane
4. PMP22 probe mix (see Table 1)
5. genomic DNA from control individual 5
- 30 6. HLPC-water for negative control
7. Hybridisation reagents

**4.2. Silanation of Whatman Anodisc 25 membranes with APS**

A 1% APS solution was prepared by filtering 3 ml of APS and subsequently adding

35 2.5 ml filtered APS to 247.5 ml HLPC-water. A 600 ml beaker was filled with the 250 ml 1% APS solution and placed on a plate shaker. A number of 50 of Anodisc 25 membranes were placed in a Teflon holder (see WO 02/072268 for further



specifications on said holder) and subsequently said holder was placed in the beaker containing the 1% APS solution. The holder was gently moved up and down to prevent that air bubbles would be enclosed between the holder and the Whatman papers. The beaker was closed with parafilm and incubated on the plate shaker at room temperature with shaking at 100 rpm for 1 hour. Subsequently, Anodisc membranes were transferred to a second beaker filled with 250 ml HLPC-water; the holder was again moved gently up and down for a couple of times and finally kept in the HPLC solution for another 3 minutes. This transfer was repeated at least one more time. The Teflon holder was then transferred 2 times to 250 ml of 96% ethanol for 3 minutes. Membranes were then transferred to an aluminium foil dish and placed for 2 hours in a vacuum oven at 120°C. After turning off the vacuum pump, membranes were allowed to cool down to room temperature after which they were stored in a dark and dry place.

**4.3. Intact genomic DNA hybridisation on Anodisc 25-APS membranes blocked with herring sperm DNA**

1 µg denatured intact genomic DNA of control individual 5 and a water control were spotted onto individual Anodisc 25-APS membranes without UV cross-linking. The DNA was allowed to dry for 10 min at room temperature. 2 ml of pre-hybridisation solution was prepared by adding together 280 µl 1M NaH<sub>2</sub>PO<sub>4</sub>, 720 µl Na<sub>2</sub>HPO<sub>4</sub>, 700 µl 20% SDS, 276 µl HLPC-water, 20 µl herring sperm DNA and 4 µl of 0.5M EDTA. The spotted Anodisc-APS membrane was pre-hybridised in 50 µl pre-hybridisation solution at 65°C for 1 hour using PamGene's flow-through system (0.2 bar pressure at 42°C for two hours at 2 cycles/min). The pre-hybridisation solution was removed and replaced with 50 µl of Cot-1 DNA solution and flown-through during 30 min at 65°C. The Cot-1 solution was subsequently removed and 50 µl of hybridisation solution comprising the PMP22 probes was added. Incubation was at 65°C during 2 hours using flow-through. It appeared that the hybridisation was difficult but not impossible.

30

**4.4. Intact genomic DNA hybridisation on Anodisc 25-APS membranes blocked with 0.5M acetic anhydride and 0.125M N,N-diisopropylethylamine**

**4.4.1. Solutions**

10 ml of blocking solution was prepared by adding together 0.47 ml 0.5M acetic anhydride, 0.2175 ml 0.125M N,N-diisopropylethylamine and 9.3 ml dichloromethane.

500 ml post-hybridisation solution 1 was prepared by diluting 25 ml 20% SSC and 25 ml 20x SDS up to 500 ml in HLPC-water.

500 ml post-hybridisation solution 2 was prepared by diluting 2.5 ml 20% SSC and 2.5 ml 20x SDS up to 500 ml in HLPC-water.

5

#### 4.4.2. Hybridisation Procedure

1 µg denatured intact genomic DNA of control individual 5 and a water control were spotted onto individual Anodisc 25-APS membranes without UV cross-linking. The DNA was allowed to dry for 10 min at room temperature.

- 10 The spotted membranes were placed with the lamination block on the washing system with vacuum (flow-through system). 20 µl of blocking solution was added to each sample and this step was repeated three more times. Samples were washed 4 times with 250 µl of 96% ethanol. During this step of the experiment, the membranes were damaged and hence no subsequent intact genomic DNA hybridisation could be performed on Anodisc-APS membranes blocked with acetic anhydride and N,N-diisopropylethylamine. The membranes could also not be washed using flow-through.
- 15

In conclusion, example 3 shows that hybridisation of intact genomic DNA immobilized onto Anodisc membranes with 2 µm pore diameter using flow-through resulted in either difficult hybridisation or damage of the membranes. Flow-through post-hybridisation washes were only possible for the water controls and not for the intact genomic DNA samples.

20

#### 25 **General Conclusion**

From Examples 1 to 4 it was surprisingly found that only matrix pore sizes above 0.45 µm in diameter allow highly efficient flow-through analysis of intact genomic DNA.

- 30 Membranes with pore sizes of 0.2 µm (Example 4) and 0.45 µm (Example 3) showed difficulties for the intact non-manipulated genomic DNA to pass through those membranes and washing off non-hybridised probes through those membranes appeared to be impossible due to the small pore sizes.

- 35 The experiments showed that pore sizes > 0.45 µm are required to assure efficient washing off of non-hybridised probes and hence highly specific hybridisations.

**Example 5: Application of the present invention in the analysis of intact genomic DNA sample for hereditary nonpolyposis colorectal cancer**

5 Colorectal cancer is one of the most common cancers in both men and women. The majority of the hereditary colorectal cancers are hereditary nonpolyposis colorectal cancer (HNPCC). HNPCC is an autosomal dominant disorder caused by mutations in mismatch repair (MMR) genes. The large majority of germline mutations detected in HNPCC families occur in the *MSH2* and *MLH1* genes. Approximately 25% of the  
10 germline mutations found in the *MSH2* and *MLH1* genes are large genomic deletions.

The present invention will allow identification of germline mutations in the MMR genes which will open the possibility of pre-symptomatic diagnosis in members of affected families. The results of a genetic screening will influence medical  
15 management of the patient or family members.

**5.1 Preparation and quality control of the probes for microarray production**

The probes of *MSH2* P1 and P2 were generated from direct amplification of human genomic DNA prepared as described in Example 1. The probes (12 control probes,  
20 16 *MSH2* probes and 19 *MLH1* probes) were generated from amplification of plasmids using the flanking vector primers PZA and PZB as described in Example 1. These plasmids were obtained from the Institute of Genetics, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK. The probes (12 control, 16 *MSH2* and 19 *MLH1*) were prepared by cloning PCR products into the  
25 EcoRV site of pZero2 (Invitrogen). The sequences of the control probes that were cloned into the plasmids can be found in Table 1. The sequences of the *MSH2* and *MLH1* that were cloned into the plasmids can be found in Tables 3 and 4 respectively. Probe quality was monitored by gel electrophoresis onto a 2% agarose gel with TAE (0.5x) at 100V for 40 min of which the results are shown in Figures 6<sub>A-C</sub>.

30 Control probes are used for normalization of microarray data and checking for PCR contamination. Human *MSH2* is located on chromosome 2p21 and has 16 exons. The 18 *MSH2* probes represent the 16 exons and two *MSH2* promoter regions (P1 and P2). Human *MLH1* is located on chromosome 3p21 and has 19 exons. The 19  
35 *MLH1* probes represent the 19 exons.

After flow-through hybridisation according to the present invention and post-hybridisation washes as described above, individual Whatman papers may be transferred in 50µl of PCR buffer (1x) in 1.5ml tubes and each boiled for 5 min. 5µl of the boiled solution is then subsequently transferred into a tube with PCR mixture comprising 5µl PCR gold buffer (10x), 3µl MgCl<sub>2</sub> (25mM), 5µl dNTPs (2.5mM), 0.125µl AmpliTaq Gold (5U/µl), 1µl PZA forward primer (10pM), 1µl PZB reverse primer (10pM), 5µl sample solution and 22.375µl HLPC water. A PCR program with following cycle order may be performed; cycle 1, 3 min at 94°C; cycles 2 to 35, 1 min at 94°C, 1 min at 60°C, 2 min at 72°C; and finally 10 min at 72°C.

10

During PCR amplification, labelled primer(s) or enzymatic generation of labelled nucleic acid may be used to generate labelled nucleic acids. The obtained labelled PCR products can then be purified using the Qiaquick purification kit and dissolved in 50 µl of EB buffer (10mM Tris-Cl, pH 8.5) for further detection using a microarray.

15

Alternatively, chemical labels can be used (Kreatech) whereby for example a platinum group forms a coordinative bond with the labelling target, firmly coupling the label to the target.

20 For analysis of HNPCC patients, 49 of 60-mer oligo's (12 control, 18 *MSH2* and 19 *MLH1*) having 40% to 50% GC content (Eurogentec) were selected for production of a microarray. The sequences of the oligo's can be found in Table 5. Upon manufacturing of a HNPCC microarray, clinical samples from patients with HNPCC can be analysed.

25

Table 1. SEQ ID Nos, names, sizes and genomic location of the control probes

SEQ ID NO.	Probe Name	Probe Size (bp)	Location	Sequence
5	8D2	522	Sub-telomeric (chrom. 19)	CTGCAGGGGCACTCAGCCCTCCTGCTGCCCTGGGTTGGCTGGCAGAGATGGCCTCACCTCTTCCAGGACACAGCAGGTGGCCACCAGCCACACACCACTCTACACAGCCACACGCGGTGTGAAGCCACAGGAAAGCATCTTAAGTACCTCTGAATCACCCCTCAGCCCCCTTAGGACCCTCTGGGTAGGGTTCGAGGACAGGTCTCTTGGGCAACCGTCTCCTCAGTCCAGGGGTGGCTTTCTACCCGCAACCAAGGAGCCTGACTAA CAAGGGATGGCAGGCGACCAATCAATGAGCCCCAGTCTGGACGGCAAGTTGACACAGTCCGCGAGCTTGAGATCCACACCGAGTAAACA GGGAGTTAACATATATGGCCCCCAGGAGAGGTGTCTGACGGTCTGAGGAGATCTGGCCCTTATATGACGTGCCATTTAAACAGGAGAG
6	8B4	451	Sub-telomeric (chrom. 18)	CTCTGGACTTTGTCATATTTTCAATAATCCTGGCTCCAAAATACTAGTCGAGGAGATCCTAGGCCACTGGGGAATTTTCTAAGGGTCTCTGAA GTAAACAATGGTAAACAGAGGTGAATTTTAGGAAAGTAAAGATGTGAACCTAGGAAAGAGATGACATAACTGAGGGGGAATGATACCCATGGGA ACAGAGAAACCTGCGTGTGAGGTGTGAGTGTGAGTGTGAGGAGACCAAGGGCTCAAGTACGCCCCCTCCGAGGGGATGGCTGTGTCGACAGATATGA CTAGAGACAACCTCTCTGGGCGGACTGCTAGAGAAACAGCAGCGCCACTGTTCGCTCTCACTGTGTGGCTGGGGAATAAAAACGCGCAGGAGGAGGA GGGACAGGAACCCAG
7	TBX5 B1	332	Sub-telomeric (chrom. 12)	gcttcttgcctcagAGCAGAACCTTGTGCGCGGACAGGGCCCTGGGGCCGACCAATGGCCGACGAGAGGGCTTTGGCCTTGGCGCACACGCGC TCTGGAGCCTGACGCAAAAGACCTGCGCTGCGATTTCGAACCCGAGAGCGCGCTCGGGGCCCCCAGCAAGTCCCCGTCGTCCCCGAGCGCCCT TCACCCAGCAGtaaggagacctgcgcgttcgggtccctttgcsagagatacaaaagtcagagtcctgcttccctgctcgggtcttctcttg
8	10B4 D	319	Sub-telomeric (chrom. 7)	CTTGCAACCTGAGTCTGCAGTCGGCTGGGTGGGGCCCTGTGCACCTGGAAGGAAGGCTGGGGCCCCGACCCAGTGTGGCTTCAGGGCCTGGTGA GGGTCACCTTTCAAGCTGCTTAGACTGACAGCCTCCCGAGGTGTGTGCTCAGGCAAGTGCCTCTGGCCGCGCACAAACGGGAAGGTGCTGTGAGGCC GCGTCTACCTGTGGCTTTAAACACTGGCTTCCACCTGCTGGCCCTCCGCGCTCCAGCCCGCGCCCTGGAG
9	TBX5 Da	290	TBX gene (chrom. 12)	cctgggtcggtgaactgaagcacgcttcgggtgcagtgcgtaacctccagactctgagccagggcctctagtaaacctctccttcatctaggtctgtg GACGGCAAGCTGAGCCCGCCATGCTGCGCGCTGTACGTGACCCAGACTCCCGCCGACCGGGCGCATTTGGATGAGGAGCTCGTCTCTCC TTCCAGAAACTCAAGCTCACCAACAACCACTGGACCCCAATTTGGGC
10	5C3	282	Sub-telomeric (chrom. 8)	CTGTGTTCCAGGTTTCGGAACCTGAGGCGGCGAGATGGGAACCTGTGCTCCCTGACCTTAAGAGCAGGAGGTGAGAGCCCTGTGGGC TGAGTAATCCTCTGAAGCACTTGTGTGGCTGGAAAGATGTGTTTTTTCAGGCTTAATCTGTATGATGATTTATCGGAAAAATGTAATTTGTCTGT GTAAACAGCAGCAGACTGGCCATTTCTGCTGGCAG
11	4A4	235	Sub-telomeric (chrom. 4)	TCCTGTGCTCCCTCCCCACACTCAGTTTCACTATGGCCAGATGCTGCTCCGATGAAGAAAGGAGGCCCATGGCAGAGCGTTCGAGGGGGCC AGGGTGGCCACACAGGCCAGGAGACCAACTCTAAACCTGATCTGACACAGGTCTAAGGGGAAGTTCATGAAGAAACAACAG

Table 1 (continued)

SEQ ID NO.	Probe Name	Probe Size (bp)	Location	Sequence
12	A2Y	222	Y-linked probe (chrom.Y)	CTGTAACTCTAAGTATCAGTGTgAAACGGGAGAAACAGTAAAGGCAACGTCAGGATAGAGTGAAGCGACCCCATGAACGCATTTCATCGTGTGGTCTCCGATCAGAGGGCGCAAGATGGCT CTAGAGATCCCAGAAATGCGAAACTCAGAGATCAGCAAGCAG
13	11G9	188	Sub-telomeric (chrom.19)	CTGCAGTGGGATGAGATGGGCTGAGGTTTGTGCCCTGTAGCCGTATGTGAACCATGGGGCAAGGTGGTCAGCGGGGGTCAGAGGTATTGTACAA GGGTCCACATAGGAATGGCAGGGTGTGAGCAG
14	TBX5A2	170	TBX gene (chrom.12)	Tcaaagtcttcacctgagccctgctctccagcgaggcgactcctggcttttgcgtccaaagaagaggtgggatatgttggaggtagtcttcacc ctggaggactgagggg
15	2D2	167	X-linked probe (chrom.X)	CCTCCCCTGCTCAGCAGCTCCTGGGATTGGAACCTGCTTCTCTCTGCAAGCCTCCTAGCCCCGGTTCTCCAGCCCTCCCCCAGACCAATCATGGG ATAGTGCCTAGG
16	D11	153	Non-human probe	CCACTACGTGAACCATCACCCCTAATCAAGTTTTTTGGGGTTCGAGGTGCCGTAAAGCAGCTAAATCGGAACCCCTAAAGGGAGAGCCCCCGATTTAGAG

**Table 2.** Oligonucleotides used for direct amplification of genomic DNA for generation of *MSH2* probes. MT, melting temperature; AT, annealing temperature; GC, GC content; P1, promoter region 1 of *MSH2* gene; P2, promoter region 2 of *MSH2* gene.

Values for size and GC-content (%) indicated are of the amplified PCR products using a pair of forward (F) and reverse (R) primers.

SEQ ID No.	exon	Primers, 5'→3'	Size (bp)	MT (°C)	AT (°C)	GC (%)
17	P1F	AGTAACGGCCGCCAGTGTGCTGGTTTTCAATCTGTCGCCAC	189	66	56	57
18	P1R	CGAGCGGCCGCCAGTGTGATGCCTGGGCAACATGGTAAAC	189	66	56	57
19	P2F	AGTAACGGCCGCCAGTGTGCTGCAATCTCCTGGGCTCAAGTG	218	68	58	45
20	P2R	CGAGCGGCCGCCAGTGTGATGACTGCCTTTATTCTGCTTAC	218	62	52	45
21	1F	AGTAACGGCCGCCAGTGTGCTGGCTTCGTGCGCTTCTTTCAG	199	68	58	66
22	1R	CGAGCGGCCGCCAGTGTGATGCCGCCCATGTACTTGATC	199	70	60	66
23	2F	AGTAACGGCCGCCAGTGTGCTGAAGTCCAGCTAATACAGTGC	202	64	54	31
24	2R	CGAGCGGCCGCCAGTGTGATGCAACTCTATACTGACGAACC	202	64	64	31
25	3F	AGTAACGGCCGCCAGTGTGCTGGGTAACAATGATATGTCAGC	208	62	52	45
26	3R	CGAGCGGCCGCCAGTGTGATGGAGAGAGCCTCAAGATTGG	208	68	58	45
27	4F	AGTAACGGCCGCCAGTGTGCTGATAGATAATTCAAAGAGGAG	199	58	48	38
28	4R	CGAGCGGCCGCCAGTGTGATGTGTACCTGATTCTCCATTTTC	199	62	52	38
29	5F	AGTAACGGCCGCCAGTGTGCTGTTAGGTTGCAGTTTCATCAC	193	62	52	38
30	5R	CGAGCGGCCGCCAGTGTGATGAAAAGGTTAAGGGCTCTGAC	193	64	54	38
31	6F	AGTAACGGCCGCCAGTGTGCTGGTTTTCACTAATGAGCTTGC	221	62	52	39
32	6R	CGAGCGGCCGCCAGTGTGATGCTATTCTGTTCTTATCCATG	221	60	50	39
33	7F	AGTAACGGCCGCCAGTGTGCTGGTAGAAGATGCAGAATTGAG	208	62	60	39
34	7R	CGAGCGGCCGCCAGTGTGATGCAGAGCCTGTATAACATTAG	208	62	60	39
35	8F	AGTAACGGCCGCCAGTGTGCTGTTCTTTTAGGAAAACACCAG	228	60	50	31
36	8R	CGAGCGGCCGCCAGTGTGATGCTTTCTTAAAGTGGCCTTTG	228	62	52	31
37	9F	AGTAACGGCCGCCAGTGTGCTGCTTTGTTCTGTTTGCAGGTG	208	64	54	36
38	9R	CGAGCGGCCGCCAGTGTGATGCAACCTCCAATGACCCATTC	208	66	56	36
39	10F	AGTAACGGCCGCCAGTGTGCTGTTGTTTATCAAGGCTTGGAC	196	62	52	35
40	10R	CGAGCGGCCGCCAGTGTGATGATTTAACACCATTCTTCTGG	196	60	52	35
41	11F	AGTAACGGCCGCCAGTGTGCTGCTGTTATTTTCGATTTCAGC	195	62	52	29
42	11R	CGAGCGGCCGCCAGTGTGATGGACATTCAGAACATTATTAG	195	58	48	29

Table 2 (continued)

SEQ ID No.	exon	Primers, 5'→3'	Size (bp)	MT (°C)	AT (°C)	GC (%)
43	12F	AGTAACGGCCGCCAGTGTGCTGACTCAATGATGTGTTAGCTC	203	62	52	41
44	12R	CGAGCGGCCGCCAGTGTGATGTTTCATCTTGAACCTCAACAC	203	60	50	41
45	13F	AGTAACGGCCGCCAGTGTGCTGTCGACAACTGGGGTGATAG	192	66	56	49
46	13R	CGAGCGGCCGCCAGTGTGATGTTTCAGCCATGAACGTGGAG	192	66	56	49
47	14F	AGTAACGGCCGCCAGTGTGCTGGGAACTTCTACCTACGATGG	207	66	56	40
48	14R	CGAGCGGCCGCCAGTGTGATGGTGGTGAGTGCTGTGACATG	207	68	58	40
49	15F	AGTAACGGCCGCCAGTGTGCTGTCTTATAGGTGTCTGTGATC	211	62	52	40
50	15R	CGAGCGGCCGCCAGTGTGATGCACTTCTTTGCTGCTGGTTC	211	66	56	40
51	16F	AGTAACGGCCGCCAGTGTGCTGCCAAGGTGAAACAAATGCCC	208	66	56	31
52	16R	CGAGCGGCCGCCAGTGTGATGACCTTCATTCCATTACTGGG	208	64	54	31



Table 3. MSH2 exon probe sequences, gene ID, name and size

SEQ ID No.	Probe Name	Size	Exon	Sequence
53	P1	198	Promo ter	gttttcaatc tgtcgcccaac gctggagtagc agtgggcaaa tttacggctg caccgcagcc tcgacctccc gggcteaggt gatccttttcg cctcagccct gctaatactt gggatcacag acgtgggttt taccatgttg ccag
54	P2		Promo ter	caatctcctg ggtcaagt atccgcccaac ctgggcctcc caaatgtctg ggattacagg cgtgagctac cgcgcctgc cacaacgca tatcttctaa cgtaccattt catttacttg ctatattcat tatctgaatt ttctcatatt agaattgaag cagaataaag gcagt
55	S1	338	1	gcgggaaca gcttagtggg tgtggggctg cgcattttct tcaaccaggga ggtgagaggg tttcgcacatg gcggtgcagc cgaaggagac gctgcagttg gagagcggc cccaggtcgg ctctgtgctc ttctttcagg gcattccgga gaagccgacc accacagtc gccttttcga cccggccgac ttctatatacgg cgcacggcga ggacggcctg ctggccgccc gggagggtgt caagaccag ggggtgatca agtacatggg gccggcag
56	S2	220	2	ggagcaaga atctgcagag tgtgtgctt agtaaatga atttgatc ttttgtaaa gatcttcttc tggttcgtca gtatagagtt gaagtttata agaatagagc tggaataaag gcatacagga gaatgattgg tatttggcat atraag
57	S3	332	3	cctggcaatc tctctcagtt tgaagataft ctctttggta acaatgatat gtcagcttcc attggtgttg tgggtgttaa aatgtccgca gttgatggcc agacacaggt tggagttggg tatgtggatt ccatacacag gaactagga ctgtgtgaat tcctgtatga tgatcagttc tccaatcttg aggcctctct cttccagatt ggaccaaagg aatgtgtttt acccgagga gagactgctg gagacaatgg gaaactgaga cag
58	S4	220	4	aatagataa ttcaaagagg aggaattctg atcacagaaa gaaaaaagc tgaacttttc acaaaagaca tttatcagga cctcaacccg ttgttgaaag gcaaaaagg agacagatg aatagtgctg tatggccaga atggagaaat caggtacatg g
59	S5	209	5	gttgcaattt catcactgtc tgcggtaatc aagtttttag aactcttatc agatgatttc aactttggac agtttgaact gactactttt gacttcagcc agtatatgaa attggataft gcacagtca gagcccttaa ctttttcag
60	S6	195	6	sggttctgtt gaagatacca ctgctctca gtctctgctt gccttggct ataatgttaa aacccctcaa ggacaagac ttgttaacca gtggatttag cagcctctca tggataagaa cagaatagag gagag
61	S7	263	7	cagattgaat ttatgggaag cttttgtaga agatgcagaa ttgaggcaga cttttacaga agatttactt cgtcgattcc cagatcttaa ccgacttgcc aagaagttc aaagacaagc agcaactta caagattgtt accgactcta tcagggtata aatcaactac ctaatgttata caggctctg aaaaacatga agg

Table 3 (continued)

SEQ ID No.	Probe Name	Size	Exon	Sequence
62	S8	176	8	GGAACACACC AGAATTAATT GTTGGCAGTT TTTGTGACTC CTCTTACTGA TCTTCGTTCT GACTTCTCCA AGTTTCAGGA AATGATAGAA ACAACTTTAG ATATGGATCA Ggtatgc
63	S9	188	9	GcagGTGGAA AACCATGAAT TCCTTGTAAG ACCTTCATTT GATCCTAATC TCAGTGAATT AAGAGAAATA ATGAATCACT TGGAAAAGAA GATGCAGTCA ACATTAATAA GTGCAGCCAG AGATCTTg
64	S10	249	10	gtttatcaag GGCTTGGACC CTGGCAACA GATTAACTG GATTCAGTG CACAGTTGG ATATTACTTT CGTGTAACTT GTAAGGAAGA AAAAGTCCTT CGTAACAATA AAACTTTAG TACTGTAGAT ATCCAGAAGA ATGGTGTAA ATTTACCAAC Aggtttgcaa gtcgttatta tatttttaac c
65	S11	159	11	GAAGCCAGG ATGCCATTGT TAAAGAAATT GTCAATATTT CTTCAGgtaa acttaataga actaataatg ttctgaatgt cacctggcctt ttggtaacag
66	S12	307	12	GGCTATGTAG AACCAATGCA GAACTCAAT GATGTGTAG CTGAGCTAGA TGCTTTGTC AGCTTTGTC ACGTGTCAA TGGAGCACCT GTTCCATATG TACGACCAGC CATTTGGAG AAGGACAAAG GAAGAATAT ATTAAGAACA TCCAGGCATG CTGTGTGTGA AGTTCAAGAT GAAATGTCA TTAATTCCTAA TGACGTATAC TTTGAAAAG ATAAACAGAT GTTCCACATC ATTACTGg
67	S13	266	13	GGCCCCAATA TGGGAGGTAA ATCAACATAT ATTCGACAAA CTGGGGTGAT AGTACTCATG GCCCAATTG GGTGTTTTGT GCCATGTGAG TCAGCAGAAG TGTCCATTGT GGAATGCAATC TTAGCCCGAG TAGGGGCTGG TGACAGTCAA TTGAAGAGAG TCTCCACGTT CATGGCTGAA ATGTTGGAA CTGCTTCTAT CCTCAGg
68	S14	315	14	cagGTCTGCA ACCAAGATT CATTAATAT CATAGATGAA TTGGGAAGAG GAACTTCTAC CTACGATGGA TTTGGGTAG CATGGGCTAT ATCAGAATAC ATTGCAACAA AGATTGGTGC TTTTGTGATG TTTGCAACCC ATTTTCATGA ACTTACTGCC TTGGCCATC AGATACCAAC TGTTAATAAT CTACATGTCA CAGCACTCAC CACTGAAGAG ACCTTAACATA TGCTTTATCA GGTGAAGAAA Ggtatg
69	S15	237	15	SGTGTCTGTG ATCAAAGTTT TGGGATTCAT GTTGCAGAGC TTGCTAATTT CCTAAGCAT GTAATAGAGT GTGCTAAACA GAAAGCCCTG GAAC TTGAGG AGTTTCAGTA TATGGAGAA TCGCAAGGAT ATGATATCAT GGAACCAAGCA GCAAGAAGT GCTATCTGGA AAGAGAGg
70	S16	153	16	ctcatgggac attcacatgt gtttcagCAA GGTGAAAAA TTATTCAGGA GTTCTGTGCC AAGGTGAAC AAATGCCCTT TACTGAAATG TCAG

Table 4. *MLH1* exon probe sequences, gene ID, name and size

SEQ ID No.	Probe Name	Size	Exon	Sequence
71	L1	243	1	CTTCCGTGA GCATCTAGAC GTTTCCTTGG CTCCTCTGGC GCCAAATGT CGTTCGTGGC AGGGTTATT CGGCGGCTGG ACAGACACAGT GGTGAACCGC ATCGCGCGG GGBAAGTTAT CCAGCGGCCA GCTAATGCTA TCAAAGAGAT GATTGAGAAC Tggtaacggag ggagtcgagc cggg
72	L2	276	2	ctcatattaa aatatgtaca ttagagtagt tgcagactga taaattattt tctgtttgat ttgccagttt agatgcAAAA TCCACAAGTA TTCAAGTGAT TGTAAAGAG GGAGGCTGA AGTGAATTCA GATCCAAGAC AATGGCACCG GGATCAGGgt aagtaaaacc tcaaagtagc aggatgttg tgcgcttcat ggaagagtoa ggacctttct c
73	L3	224	3	gagatttggg aaaaagagta acatgattat ttactcatct ttctgtatc taacagAAAG AAGATCTGGA TATTGTATGT GAAAGGTTCA CTACTAGTAA ACTGCAGTCC TTTGAGGATT TAGCCAGTAT TTCTACCTAT GGCTTCGAG GTGAGgtaag ctgag
74	L4	171	4	cttttcttcc tttagGCTTGG GCCAGCATAA GCCATGTGGC TCATGTTACT ATTACAACGA AACACAGCTGA TGGAAAGTGT GCATACAGgt atagtgtga cttctttttac tc
75	L5	288	5	ttgatatgat ttctctttt ccccttggga ttagtatcta tctctctact ggataattaat ttgttatatt ttctcattag AGCAAGTTAC TCAGATGGAA AACTGAAAGC CCCTCCTAAA CCATGTGCTG GCAATCAAGG GACCCAGATC A
76	L6	257	6	gggttttatt ttcaagtact tctatgaatt tacaagaaaa atcaatcttc tgttcagGTG GAGGACCTTT TTTACAACAT AGCCACGAGG AGAAAAGCTT TAAAAATCC AAGTGAAGAA TATGGAAAA TTTTGGAAgt TGTGGCAGg tacagtccaa aatctgggag tgggtctctg agattttgtca tcaaa
77	L7	167	7	ggctctgaca tctagtgtgt gtttttggca actcttttct tactcttttg ttttctttt ccagGTATTC AGTACACAAT GCAGGCATTA GTTCTCAGT TAAAAAag
78	L8	149	8	CAAGGAGAGA CAGTAGCTGA TGTTAGGACA CTACCAATG CCTCAACCGT GGACAATATT CGCTCCGTCT TTGGAATGC TGTTAGTCGg
79	L9	184	9	GAGAACTGAT AGAAATTGGA TGTGAGGATA AAACCTTAGC CTTCAAAATG AATGGTTACA TATCCAATGC AAATCTACTCA GTGAAGAAgt GCATCTTCTT ACTCTTCATC AACGgtaagt taaaa
80	L10	327	10	ttattgttta GATCGTCTGG TAGAATCAAC TTTCTTGAGA AAAGCCATAG AAACAGTGTA TGCAGCCTAT TTGCCAAAA ACACACACCC ATTCTGTAC CTCAGgtaat gtageaccaaa actctctaac caagactcac aggaacacaga tgtctatca ggctctcttc ttgaaagag atgagcatgc taatagtaca atcagagtga atcccataca ccactggcaa aggatgttc tgccttctt tacaggtaca aggcacag
81	L11	274	11	cctgacagTT TAGAAATCAG TCCCCAGAAT GTGGATGTTA ATGTGCACCC CACAAGCAT GAATTCACt TCCTGCACGA GGAGAGCATC CTGGAGCGG TGCAGGAGCA CATCGAGAGC AAGCTCCTGG GCTCCAATTC CTCCAGGATG TACTTCACCC Aggtcagggc gcttctctac cagctacttc tctctggggc ctttgaaatg tgcccg

Table 4 (continued)

SEQ ID No.	Probe Name	Size	Exon	Sequence
82	L12	359	12	CCTCTGGGGA GATGGTTAAAT CCACAACAAGT CTGACCTCGT CTTCTACTTC TGGAAGTAGT GATAAGGTCT ATGCCACCA GATGGTTCGT ACAGATTCCC GGGAAACAGAA GCTTGATGCA TTCTTGACG CTTCTGACAA ACCCTGTCC AGTCAGCCCC AGGCCATTGT CACAGAGGAT AAGACAGATA TTCTTAGTGG CAGGGCTAGG CAGCAAGATG AGAGATGCT TGAACGCCA GCCCTGCTG AAGTGGCTGC CAAAATCAG AGCTTGGAGG GGGATACAC AAGGGGA
83	L13	257	13	cctttcttc attgcagaaa GAGACATCGG GAAGATTCTG ATGTGMAAT GTTGAAGAT GATTCCGAA AGGAAATGAC TGCAGCTTGT ACCCCCCGGA GAAGATCAT TAACCTCACT AGTGTCTTGA GTCTCCAGGA AGAAATTAAT GAGCAGGAC ATGAGGgtac gtaaacgtg tgccctgcct gggatgcata ggg
84	L14	205	14	gcagTCTCC GGGAGATGTT GCATAACCAC TCCTTCGTGG GCTGTGTGAA TCCTCAGTGG GCCTTGGCAC AGCATCAAAC CAAGTTATAC CTTCTCAACA CCACCAAGCT TAGataaatc agctgagtg gtgtaacaag cagagct
85	L15	199	15	cagTGAAGAA CTGTTCTACC AGATACTCAT TTATGATTTT GCCAATTTTG GTGTCTCAG GTTATCGta agtttagatc ctttccactt ctgacatttc aactgaccgc ccgcaaca gtagctctcc actaaata
86	L16	303	16	cctagGAGCC AGCACCCTC TTGACCTTG CCATGCTTGC CTTAGATAGT CCAGAGAGTG GCTGGACAGA GGAAGATGTT CCCAAAGAAG GACTTGTCTGA ATACATTGTT GAGTTTCTGA AGAAGAAGCC TGAGATGCTT GCAGACTATT TCTCTTTGGA AATTGATGAG gtgtgacagc cattcttata cttctgttgt attctccaaa taaaatttcc agccgggtgc attggc
87	L17	311	17	gttcccttgt cctttttcct gcaagcagGA AGGAAACCTG ATTGGATTAC CCTTCTGAT TGACAACTAT GTGCCCCCTT TGGAGGGACT GCCTATCTTC ATCTCTCGAC TAGCCACTGA Ggtcagtgat caagcagata ctaagcattt cgttacatgc atgtgtgctg gaggaa
88	L18	229	18	gaggtattga atttcttttg accagGTGAA TTGGGACGAA GAAAGGAAAT GTTTGAAAG CCTCAGTAAA GAATGCGCTA TGTCTATTTC CATCCGGAAG CAGTACATAI CTGAGGAGTC GACCTCTAG GCCAGCAGt acagtggta tgacactggc accccaggac
89	L19	322	19	ccagAGTGA AGTGCCTGG CTCCATTCC AAACCTCCTG GAAGTGGAC TGTGGAACA CATTGTCTA TAAAGCCTT GCGCTCACA CATCTGCC TCCTAAACA TTTCACAGA AGATGGAAA TATCCTGCA GCTTGCTAA CCTGCTGA TCTATACAA AGTCTTTGA GAGGTGTTA AATATGGTT ATTTATGCA CTGTGGGAT GTGTTCTTC TTCTCTGT ATTCCGATA CAAAGTGT GTATCAAAG TGTGATATA CAAAGTGA CC

Table 5. 60-mer Oligonucleotides of control, *MSH2* and *MLH1* for production of HNPCC microarrays

SEQ ID No.	Gene ID	Oligo Name	GC-content	Exon	Sequence
90	sub-telomeric	8D2-1	58%	Control	TGGCTTTCTACTCCCGCAACAGGAGCCTGACTAAACAGGGATGGCAGGCCACACATC
91	sub-telomeric	8B4-2	51%	Control	GAAATGATACCCATGGGAACAGAGAAACCTGCGTGTGAGGTGTGAGGTGTGAGGAGACCAG
92	sub-telomeric	TBX5B1-3	56%	Control	CCAGCAGGTAAGGAGACCTCGCGCTTCGGGTCCTTTGCGAGAGATCAAAGTCAGAGTCTG
93	sub-telomeric	10B4D-4	61%	Control	CTGGTGAGGTCACCTTTCAGGCTGCTTAGACCTGCAGCCTCCCCAGGTGGTGTGCTCAGG
94	TBX gene	TBX5Da-5	55%	Control	CTGGTGAGGTCACCTTTCAGGCTGCTTAGACCTGCAGCCTCCCCAGGTGGTGTGCTCAGG
95	sub-telomeric	5C3-6	55%	Control	GACCTTAAGAGCAGGAGGTGAGAACCCCTGTGGGCTGAGTAATCCTCTGAAGCACCTTGC
96	sub-telomeric	4A4-7	55%	Control	CACAGGCCAGGAGACCAACCTCTAACCTGTGATCTGACACAGGTCTAAGGGAGGTTCATG
97	Y-linked	A2Y-8	48%	Control	GAAAAACAGTAAAGGCAACGTCCAGGATAGAGTGAAGCGACCCATGAACGCAATTCATCGTG
98	sub-telomeric	11G9-9	56%	Control	GTAGCCGTATGTGAACCATGGGGCAAGGTGTTGTCAGCGGGGTCAGAGGTATTGTACAAGG
99	TBX gene	TBX5A2-10	53%	Control	CACCTCTGGCTTTTGGCTCCAAAGAGAGGTGGGATAGTTGGAGGTGAGTTTCACCCCTG
100	X-linked	2D2-11	56%	Control	CTCAGCACCTCTGGGATTTGGAAACCTCGTTCTCTCTGCAAGCCTCCTAGCCCCGGTTCT
101	Non-human	D11-12	46%	Control	CTAATCAAGTTTTTTGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCCTAAAGGGAG
102	MSH2	MSH2-p1	58%	Promoter	GTTTTCAATCTGTCGCCACCGTGGAGTGCACTGGGCACAAATTTACGGCTG CACCGCAGCC
103	MSH2	MSH2-p2	51%	Promoter	CAAAATTGCTGGATTACAGCGGTGAGTACCGCGCCCTGCCCACAAACGCATATCTCTAA
104	MSH2	MSH2-1	60%	1	CTTCGTGCGCTTCTTTTCAGGGCATGCCGGAGAAGCCGACCCACACAGTGGCGCTTTTCGA
105	MSH2	MSH2-2	31%	2	CTGCAGAGTGTGTGCTTAGTAAATGAATTTTGAATCTTTTGTAAAGATCTTCTCTG

Table 5 (continued)

SEQ ID No.	Gene ID	Oligo Name	GC-content	Exon	Sequence
106	MSH2	MSH2-3	46%	3	GACAGGTTGGAGTTGGGTATGTGGATTCCATACAGAGGAACTAGGACTGTGTGAATTCC
107	MSH2	MSH2-4	48%	4	GGACCTCAACCGGTTGTTGAAGGCAAAAAGGAGAGCAGATGAATAGTGTGTTATTTGCC
108	MSH2	MSH2-5	41%	5	GAACTGACTACTTTTGTGACTTCAGCCAGTATATGAAATTGGATATTGCAGCAGTCAGAGCC
109	MSH2	MSH2-6	48%	6	CTGTTGAAGATACCACTGGCTCTCAGTCTCTGGCTGCCCTTGTCTGAATAAGTGTAAACCCC
110	MSH2	MSH2-7	45%	7	CTTCGTGATTTCCAGATCTTAAACCGACTTGGCCAAGAGTTTCAAAGACAAGCAGCAAAAC
111	MSH2	MSH2-8	41%	8	GCAGTTTTTGTGACTCCTCTTACTGATCTTCGTTCTGACTTCTCCAAGTTTTCAGGAAATG
112	MSH2	MSH2-9	35%	9	GGTGGAAAACCATGAATTCTTGTGTAAACCTTTCATTTGATTCCTTAATCTCAGTGAATTAAG
113	MSH2	MSH2-10	43%	10	GCCTGGACCCTGGCAAAACAGATTAAACTGGATTCAGTGCACAGTTTGGATATTACTTTTC
114	MSH2	MSH2-11	33%	11	GACTTCTTTTAAATGAAGAGTATACCAAAAATAAAAAACAGAAATATGAAGAAGCCCAAGGATGC
115	MSH2	MSH2-12	48%	12	CTCAGCTAGATGCTGTGTGTCAGCTTTGCTCAGTGTCAAATGGAGCACCTGTTTCCATATG
116	MSH2	MSH2-13	51%	13	GGTGTTTTGTGCCATGTGAGTCAGCAGAAAGTGTCCATTGTGGGACTGTCATCTTTAGCCCGAG
117	MSH2	MSH2-14	41%	14	GAACTTACTGCCCTGGCCCAATCAGATACCAACTGTGTTAATAATCTACATGTACAGCACTC
118	MSH2	MSH2-15	41%	15	GGGATTTCATGTTGCAGAGCTTGTCTAATTTCCCTAAGCATGTATAAGAGTGTGCTTAAACAG
119	MSH2	MSH2-16	33%	16	GCCCTTTACTGAATGTGAGAAGAAAACATCAATAAAGTTAAAAACAGCTAAAGCTGA
120	MLH1	MLH1-1	56%	1	CAAAAATGTCGTTCTGTGGCAGGGGTTATTTCGGCGGCTGGACGACAGTGGTGAACCCGCAT
121	MLH1	MLH1-2	41%	2	CAAGTGATTGTTAAAGAGGGAGGCCCTGAAGTTGATTTCAGATCCAAAGACAATGGCACCGGG

Table 5 (continued)

SEQ ID No.	Gene ID	Oligo Name	GC%	Exon	Sequence
122	MLH1	MLH1-3	41%	3	GGTTCACTACTAGTAAACTGCAGTCCTTTGAGGATTTAGCCAGTATTCTACCTATGGCT
123	MLH1	MLH1-4	45%	4	GGCCAGCATAAAGCCATGTGGCTCATGTACTATTACAAACGAAAAACAGCTGATGGAAAGTG
124	MLH1	MLH1-5	48%	5	GCAAAGTTACTCAGATGGAAAACTGAAAAGCCCCCTCTCTAAACCATGTGCTGGCAATCAAGGG
125	MLH1	MLH1-6	38%	6	GGAGGACCTTTTACACATAGCCACGAGGAGAAAAAGCTTTAAAAAATCCAAGTGAAGA
126	MLH1	MLH1-7	35%	7	CTCTTTTGTGTTTTCTTTCCAGGTATTTCAGTACACATGTCAGGCATTAGTTTCTCAGTTA
127	MLH1	MLH1-8	50%	8	GACAGTAGCTGATGTAGGACACTACCCAATGCCTCAACCGTGGACAATATTGCTCCGT
128	MLH1	MLH1-9	40%	9	CCCTAGCCTTCAAATGAATGGTTACATATCCAATGCAAACTACTCAGTGAAGAGTGCA
129	MLH1	MLH1-10	46%	10	GCCATAGAAACAGTGTATGCAGCCTATTGCCCCAAAAACACACCCCATTCCTGTACCTC
130	MLH1	MLH1-11	56%	11	CAAAGCATGAAGTTTCACTTCTCTGCAGGAGGAGCATCCTGGAGCGGGTGCAGCAGCACA
131	MLH1	MLH1-12	51%	12	GGTTGCTACAGATTCCCGGGAACAGAAAGCTTGATGTCATTTCTGCAGCCTCTGAGCAAAAC
132	MLH1	MLH1-13	45%	13	CATCGGGAAGATTCTGATGTGGAATGGTGGAAAGATGATTCCCGAAAGGAAATGACTGCA
133	MLH1	MLH1-14	55%	14	GCATAACCACTCCTTCGTGGGCTGTGTGAATCCTCAGTGGGCTTTGGCACAGCATCAAAC
134	MLH1	MLH1-15	36%	15	GAACTGTCTACCCAGATACTCATTTATGATTTTGGCCAATTTTGGTGTCTCTCAGGTTATCG
135	MLH1	MLH1-16	55%	16	GACCTTGCCATGCTTGCCCTTAGATAGTCCAGAGATGGCTGGACAGAGGAAGATGGTCCC
136	MLH1	MLH1-17	51%	17	GGGAACCTGATTTGGATTACCCCTTCTGATTGACAACTATGTGCCCCCTTTGGAGGGACTG
137	MLH1	MLH1-18	45%	18	CCTCAGTAAAGAAATGGCTATGTCTTCTATTCCATCCGGAAGCAGTACATATCTGAGGAGTC
138	MLH1	MLH1-19	50%	19	CTCCTGGAAGTGGACTGTGGAACACATTTGTCTATAAAGCCCTTGGGCTCACACATTTCTGCC